Health Care Financing

Grants and Contracts Report

National Estimates of Nonacute Hospital Utilization for 1981

PUBS
RA
407
.3
N33
1984

Health Care Financing Administration

Health Care Financing

Grants and Contracts Report

The Office of Research and Demonstrations, Health Care Financing Administration, directs more than 300 intramural and extramural research, demonstration, and evaluation projects. These projects seek alternate ways to finance, organize, and deliver health services, as well as assess the impact of Federal programs on health care costs, providers, and beneficiaries. The Health Care Financing *Grants and Contracts Report* series represents the final reports from selected extramural projects funded by the Office of Research and Demonstrations. The statements and data contained in each report are solely those of the contractor or grantee and do not express any official opinion of or endorsment by the Health Care Financing Administration.

The National Estimates of Nonacute Hospital Utilization for 1981 presents national baseline estimates of inappropriate utilization in acute hospitals prior to the implementation of the Medicare prospective payment system. A total of 1,925,000 or 5.7 percent of adult admissions and 12 percent of associated hospital days were nonacute. Of the nonacute days identified, 4.2 percent were innappropriate because the admission itself was inappropriate. Another 7.8 percent of nonacute days occurred in stays where the admission itself was justified. These estimates were based on a probability sample of 5,732 admissions and 24,051 days of care from 253 hospitals in 40 States for adult patients discharged in 1981. The Standardized Medreview Instrument (SMI)—a medical utilization review tool for assessing the appropriateness of an admission and a day of care within the stay—was developed for use in this study.

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

RA
407.3
.N33
1984
c.2

Health Care Financing

Grants and Contracts Report

National Estimates of Nonacute Hospital Utilization for 1981

U.S. Department of Health and Human Services
Health Care Financing Administration
Office of Research and Demonstrations
Baltimore, Maryland
September 1984

This report was prepared under Health Care Financing Administration (HCFA) Contract No. 500-80-0053 between HCFA and SysteMetrics, Inc. The project director was Christy Moynihan. The HCFA Project Officer for this contract was Sherry A. Terrell.

The project director expresses appreciation to Cathleen Barnes, Kelly Peinado, Nancy Jordan, Nancy Thomasco, Donald Schroeder, and Chester Yang for managerial and administrative support.

TABLE OF CONTENTS

SECTION	TITLE	PAGE
LIST OF EXHIBITS		v
I.	EXECUTIVE SUMMARY	I- 1
	A. Introduction	- 1
	B. Methods	- 2
	1. Instrument Development	- 2
	2. Sample Design	- 4
	3. Data Collection	- 5
	4. Response Bias	- 5
	5. Representativeness of the Sample	- 6
	6. Estimation Procedures	- 7
	7. Estimation of Sampling Variances	- 7
	8. Assessment of Inter-Rater Reliability	- 8
	C. Study Results	- 8
	1. The Magnitude of Nonacute Hospital Utilization	- 8
	2. Factors Associated with Nonacute Utilization	- 9
	D. Discussion	-19
	1. How much nonacute utilization is associated with inefficient scheduling and delays in discharge	-19
	2. How much nonacute utilization is associated with reimbursement mechanisms (especially insurance coverage) which encourage overutilization of hospital care	-20
	3. How much nonacute utilization is associated with differences in hospital structure and organization	I-21
	4. How much nonacute utilization is associated with area characteristics, indicative of resource availability, economic incentives, or practice style	-22

SECTION TITLE PAGE

E. Conclusions and Implications . . . . -25

F. Suggestions for Further Research . . . . -29

II. INTRODUCTION AND BACKGROUND. . . . . II- 1

A. The Magnitude of Unnecessary Hospital Utilization. - 2

1. Assessment of Inappropriate Utilization Through Comparison of Utilization Rates. . . - 2

2. Direct Assessments of Inappropriate Utilization. . . . . - 6

B. Causes of Unnecessary Hospital Use . . . . - 7

1. The Physician's Role in Inappropriate Utilization. . . . . - 8

2. The Effects of Insurance Coverage. . . . . . -10

3. Hospital Characteristics Related to Inappropriate Utilization. . . . . -13

4. The Availability and Appropriate Utilization of Alternate Care Facilities and Services. . -15

5. The Impact of Utilization Review Programs. . -19

C. Summary . . . . -26

III. DESIGN OF THE STUDY. . . . . III- 1

A. Development of the Methodology . . . . - 1

1. Overview . . . . - 1

2. Status of Utilization Review Methodologies at Project Initiation. . . . . - 4

3. Design of the Preliminary Instrument . . . . -11

4. Evaluation of the Preliminary Instrument . . -18

5. Instrument Revision Process. . . . . -24

6. Description of the Final Methodology . . . . -27

B. Design of the National Survey. . . . . -34

1. Development of the Sampling Plan . . . . . . -34

2. Data Collection. . . . . -45

SECTION TITLE PAGE

IV. DATA ANALYSIS AND METHODOLOGICAL ISSUES. . . . . . . IV- 1
  A. Description of the Sample. . . . . - 1
    1. Sample Characteristics . . . . - 1
    2. Response Bias. . . . . - 8
    3. Representativeness of the Sample . . . . . . -12
  B. Estimation and Calculation of Sampling Variance. -18
    1. Estimation Procedures . . . . -19
    2. Determination of Sampling Variance . . . . . -25
    3. Generalization of Sampling Variances . . . . -29
  C. Reliability and Validity of the National Survey Data. IV-31
    1. Assessment of Inter-Rater Reliability . . . -32
    2. Assessment of the Validity of Reviewer Evaluations. . . . . -33

V. STUDY RESULTS. V- 1
  A. National Estimates of Nonacute Utilization . . . - 1
  B. Factors Associated with Nonacute Admissions. . . - 5
    1. Descriptive Analyses . . . . - 7
    2. Multivariate Analyses. . . . . -22
    3. Summary . . . . -45
  C. Factors Associated with Nonacute Days. . . . . . -48
    1. Descriptive Analyses . . . . -48
    2. Multivariate Analyses. . . . . -71
    3. Summary . . . . -79

VI. SUMMARY AND CONCLUSIONS. . . . . VI- 1
  A. The Magnitude of Nonacute Hospital Utilization . - 1
  B. Factors Associated with Nonacute Hospital Utilization . . . . - 4
    1. Major findings of the Admission-level Analysis . . . . - 4
    2. Major findings of the Day-level Analysis . . -10
    3. Synthesis. . . . . -15

SECTION **TITLE** **PAGE**

C. Implications and Suggestions for Future Research . . . . -24

REFERENCES. . . . , ... R-1

LIST OF EXHIBITS

EXHIBIT	TITLE	PAGE
III	DESIGN OF THE STUDY.	III- 1
III-1	Pretest Analysis Inter-Observer Reliability of Most Frequently Occurring ADMISSION CRITERIA (N=555)	-20
III-2	Pretest Analysis Inter-Observer Reliability of Most Frequently Occurring DAILY CRITERIA (N=7354).	-21
III-3	Pretest Analysis Inter-Observer Reliability of Evaluations of the Appropriateness of Admissions and Days of Care . . .	-22
III-4	Decision Process for Determining the Appropriateness of Admission	-29
III-5	Reason List Used to Specify Causes for Unnecessary Admissions and Days of Care.	-31
III-6	Phases of Stay: Definitions	-33
III-7	Distribution of Hospitals in the 1980 AHA Survey by Type of Service (excluding territories and Alaska and Hawaii)	-38
IV	DATA ANALYSIS AND METHODOLOGICAL ISSUES.	IV- 1
IV-1	Unweighted Distribution of Sampled Hospitals, Weighted Distribution of Sampled Hospitals, and Distribution of Hospitals in the Population by Selected Hospital and Area Characteristics	- 2
IV-2	Unweighted Distribution of Sampled Patients by Selected Patient and Hospital Stay Characteristics .	- 4
IV-3	Distribution of Hospitals from Primary and Shadow Samples	- 9

EXHIBIT TITLE PAGE

IV-4 Comparison of Responding and Nonresponding Sampled Hospitals by Selected Hospital and Area Characteristics. IV-11

IV-5 Comparison of Estimated Admissions, Days of Care and ALOS with the 1981 American Hospital Association Annual Survey of Hospitals (AHA) and 1981 National Hospital Discharge Survey (NHDS) . . . . -14

IV-6 Comparison of Weighted Admissions by Selected Patient Characteristics with the 1981 National Hospital Discharge Survey (NHDS) . . . . -16

IV-7 Results of the Admission Evaluation (unweighted) . . -34

IV-8 Level of Service Evaluations and Level of Service Criteria for Justified Admissions (unweighted) . . . -36

IV-9 Continued Stay Evaluation and Continued Stay Criteria for Appropriate Admissions (unweighted) . . -37

IV-10 Analysis of Admissions Justified by Extension or Record Review. -40

IV-11 Analysis of Admissions Where No Criterion Met or Record Review Fails to Support . . . . -42

IV-12 Distribution by ICD-9-CM Diagnosis Code of Acute Admissions with the Same Primary Diagnosis as a Nonacute Admission . . . . -43

IV-13 Summary Results of the Review of Abstractor Evaluations. . . . . -48

IV-14 Distribution of Patient Stays by Number of Days of General Observation. . . . . -51

V STUDY RESULTS. . . . . V- 1

V-1 National Estimates of the Rate of Nonacute Hospital Admissions . . . . - 2

V-2 National Estimates of the Rate of Nonacute Hospital Days . . . . - 3

V-3 Percentage of Nonacute Admissions by Patient and Stay Factors . . . . - 8

EXHIBIT	TITLE	PAGE
V-4	Percentage of Nonacute Admissions by Hospital Characteristics.	-16
V-5	Percentage of Nonacute Admissions by Area Characteristics.	-18
V-6	Reviewer Reasons for Nonacute Admissions (Includes Only Admissions Judged Nonacute by Reviewer.). . . .	-21
V-7	Variables Used in the Patient-level Models	-23
V-8A	Patient-level Admission Analysis: Patient Characteristics. Stepwise Logit Regression Model .	-26
V-8B	Patient-level Admission Analysis: Hospital Characteristics. Stepwise Logit Regression Model. .	-28
V-8C	Patient-level Admission Analysis: Area Characteristics. Stepwise Logit Regression Model. .	-29
V-9	Patient-level Admission Analysis: Patient, Hospital, and Area Characteristics. Logit Regression Model.	-31
V-10A	Hospital and Area Variables Selected for the Hospital-level Model	V-38
V-10B	Hospital-level Intercorrelations Among Hospital and Area Variables.	-40
V-11A	Hospital-level Admission Analysis: Hospital and Area Characteristics-GLM Regression Model	-43
V-11B	Hospital-level Regression Analysis: Hospital and Area Characteristics-GLM Regression Model, Alternate Specification.	-46
V-12	Percentage of Nonacute Days by Patient and Stay Factors (Excluding All Days in Nonacute Admissions).	-49
V-13	Percentage Distribution of Patients by Length of Stay and Number of Nonacute Days in Stay (Acute Admissions Only)	-55
V-14	Percentage Distribution of Nonacute Days by Length of Stay and Day in Stay.	-56
V-15	Percentage Distribution of Hospital Stays by Day of the Week for Admission and Discharge	-60

EXHIBIT TITLE PAGE

V-16 Percentage Distribution of Admission by the Day of the Week and Number of Days Prior to the Primary Procedure. . . . . -62

V-17 Percentage of Nonacute Days by Hospital Characteristics (Excluding All Days in Nonacute Admissions). . . . . -65

V-18 Percentage of Nonacute Days by Area Characteristics (Excluding All Days in Nonacute Admissions). . . . . -66

V-19 Distribution of Reasons for Nonacute Days by the Phase of the Stay. . . . . -68

V-20A Patient-level Day Analysis: Patient Characteristics Stepwise Logit Regression Model. . . . . -72

V-20B Patient-level Day Analysis: Hospital Characteristics Stepwise Logit Regression Model. V-74

V-20C Patient-level Day Analysis: Area Characteristics Stepwise Logit Regression Model. . . . . -76

V-21 Patient-level Day Analysis: Patient, Hospital, and Area Characteristics. Logit Regression Model. . . . -77

V-22 Hospital-level Day Analysis: Hospital and Area Characteristics. GLM Regression Model . . . . -78

I. EXECUTIVE SUMMARY

I. EXECUTIVE SUMMARY

A. INTRODUCTION

Wide variation in hospital utilization rates that cannot be attributed to demographic factors, differences in severity of illness, or quality of care have been found across various geographic and organizational settings (Gornick, 1982; Bunker, 1970; Luft, 1978; Roemer and Shonick, 1973; Wennberg and Gittelsohn, 1973, 1975; Nobrega et al, 1982). Thus, indirect evidence supports the hypothesis that a non-trivial portion of acute hospital care may be unnecessary and inappropriate. Wide variations in the levels of inappropriate use, however, have been found by studies using a variety of evaluative methods to measure unnecessary hospital utilization directly (Gertman and Restuccia, 1981; Zimmer, 1967; Wenkert, Hill, and Berg, 1969; Trussell, Morehead, and Erlich, 1962). Accordingly, there has been much controversy over the magnitude of inappropriate hospital utilization and its causes.

The lack of a standard, uniformly accepted methodology for measuring the levels and causes of inappropriate utilization has hampered the efforts to resolve theoretical and practical questions. Additionally, the methods that have been used have been applied to small and/or self selected samples of respondents and are not generalizable to the nation as a whole. The purpose of the present study, therefore, was to (1) develop and rigorously test the reliability and validity of an instrument to measure levels and likely causes of inappropriate hospital utilization, and (2) apply the methodology to a nationally representative sample of hospitalizations to make reliable estimates of the rate of inappropriate utilization and to identify the factors associated with it.

Given that some of this utilization may be outside the control of hospitals or physicians, the term "nonacute" has been used in the place of "inappropriate" since, regardless of cause or considerations of social versus medical "appropriateness," the goal of the study was to identify instances where acute care hospitals are being used for other than the delivery of acute-level services.

This study is the first national assessment of nonacute hospital utilization, defined as medically-unwarranted admissions and days of care in general acute-care hospitals for the total adult U.S. population. Initiated in July 1980, the study developed a state-of-the-art methodology to measure inappropriate hospital utilization through the synthesis of existing utilization review techniques. This methodology, termed the Standardized Medreview Instrument (SMI), was extensively tested and refined in a field test conducted in 15 hospitals. Following this, the technique was applied in a national probability sample to estimate the magnitude of inappropriate utilization.

A three stage cluster sample was utilized in which medical records were retrospectively abstracted in 253 hospitals for 5,732 admissions/discharges and 24,051 days of care for 1981. After analysis of sample representativeness and abstractor reliability, ratio estimators of nonacute hospital utilization were calculated for inappropriate admissions and days of care. These were used in a variety of descriptive and multivariate analyses to identify factors associated with nonacute hospital utilization.

B. METHODS

1. Instrument Development

The philosophical approach taken by the study was to define the appropriateness or necessity of hospital utilization in terms of medical necessity alone, i.e., by whether the medical circumstances of the patient's condition require the use of hospital level resources. In the evaluation of appropriateness, the necessity of both the admission and of each day of care were judged separately. The basic methodology adopted for the study employed a retrospective review of the medical record. In addition, a criterion-based approach, rather than one based on subjective clinical judgment, was considered essential to ensure the reliability and validity of the evaluations. The appropriateness of elective surgery was not evaluated in the study.

Of the criterion-based methodologies identified in the background research, two were considered applicable for use in the present study: The Appropriateness Evaluation Protocol (AEP) developed at Boston University and Interqual's Intensity of Service, Severity of Illness and Discharge Screens (ISD). These together formed the basis of the Standardized Medreview Instrument (SMI) developed for the current study. The study adopted the pencil-and-paper application process of the AEP and included its provision for reviewer overrides of the criteria. Also included was a modified version of The AEP's reason list for the assignment of responsibility for a nonacute day. The criteria for the instrument were taken from the 13 body system specific and one generic criteria sets of the ISD. These sets contained over 200 criteria for assessing the appropriateness of the admission and subsequent days of care. In addition, modifications that had been made to these criteria sets were collected from over 50 PSROs that had been using the ISD for concurrent and retrospective review.

These preliminary criteria sets were evaluated and revised by the Medical Consultant Panel, and the modifications made by PSROs were integrated into the sets. Following this, the instrument was field tested in 15 hospitals on the east and west coasts to assess its reliability and the feasibility of applying it on a national scale. Both the reliability of abstracting criteria and the reliability of the final evaluation of appropriateness were assessed. In addition, various training and application strategies were assessed. In particular, the performance of MDs, RNs and professional medical records personnel were compared, as were mail versus face-to-face training strategies.

High levels of agreement, ranging from .77 to .89, were found for the abstracting of admission criteria. Lower levels, averaging around .55, were found for the daily criteria, a fact which was attributed to the cumbersomeness of the body system specific approach. The reliability of judgments of appropriateness were generally high (>.80) for both admissions and days. Face training produced higher reliability coefficients than mail training, but the levels for mail training were adequate, ranging from .69 to .84. Accredited Record Technicians (ARTs) and Registered Record Administrators (RRAs) were more reliable abstractors than RNs, with scores ranging from .83 to .91.

Based on the information from this field test, the pre-test instrument was substantially streamlined, and the individual body system specific criteria of the ISD collapsed into one generic set of criteria. The final instrument included 117 admission criteria, as well as 56 criteria for day review.

2. Sample Design

The national study used a three stage cluster sample: The first stage or primary sampling units (PSUs) were acute-care, short stay hospitals, the second stage consisted of records within each sampled hospital, and the third stage consisted of days within a stay. Only nonfederal, short-term, general hospitals were included in the sample. Pediatric patients and pediatric hospitals were excluded, as were routine obstetrical cases (normal delivery). In addition, no hospitals were sampled in the U.S. territories or in Alaska or Hawaii.

The basic stratification variables used in the design were region, ownership and bedsize. Proprietary hospitals were oversampled with the goal of increasing their expected numbers in the sample from 22 to 45. Within the region, ownership, and bedsize strata, 240 individual strata with approximately equal numbers of admissions per strata were devised with the goal of sampling one hospital per stratum. In addition, 15 strata representing HMO-owned hospitals were created for the sample. One hospital was selected from each stratum with probability proportional to size. In selecting hospitals from strata, a controlled selection technique was used to maximize the representativeness of sampled hospitals in terms of occupancy rate, urban/rural location, and teaching status. In this process, 10 strata contributed two, rather than one, hospitals to the sample.

Records were sampled only for patients discharged in 1981. All records had an equal probability of being selected and there was no disproportionate sampling by diagnosis or body system. Records were selected within the hospital using a systematic sampling procedure. Once a record was included in the sample, the abstractor reviewed up to seven days within the stay. Each day within a stay had an equal probability of selection.

3. Data Collection

The project was approved by the American Hospital Association (AHA) which provided a letter of endorsement for inclusion in the materials sent to hospitals requesting their participation. This was considered a critical factor in the study's successful recruitment of hospitals. Data collection for the study was initiated August 14, 1982, and continued through November 30, 1982. Recruitment and training were carried out by mail. In addition, because of their greater reliability in performing the abstracting in the field test, medical records professionals (ARTs and RRAs) were used as the data collectors rather than RNs or MDs. A total of 78 medical records abstractors travelled to or within 41 states to collect the data. Extensive quality control procedures were used to insure the reliability and accuracy of the data collected.

The final data set was comprised of 5,732 admissions and 24,051 days of care sampled from 253 hospitals.

4. Response Bias

Five recruitment waves were required to recruit the 253 hospitals in the sample. Of these, 167 hospitals responded to the initial request, for a primary response rate of 66%. Only three of the 15 HMO-owned hospitals agreed to participate, however.

Replacement hospitals were recruited from the same stratum as the non-responding hospitals. Thus, the final sample distribution necessarily resembles the distribution of the primary sample on bedsize, region, ownership and the other stratification variables. A comparison of the primary respondents with the nonrespondents in terms of bedsize, total admissions, total days of care, occupancy rate, average length of stay, region, teaching status, SMSA size, and ownership was made. There were no differences between primary respondents and nonrespondents on any of the utilization measures (admissions, days of care or ALOS) or on occupancy rate. Respondents were significantly more likely to come from areas of moderate population than nonrespondents. There was a tendency for nonrespondents to be proprietary hospitals and nonteaching hospitals, but these trends were not statistically significant.

5. Representativeness of the Sample

Two sources were available to assess the representativeness of the national sample: (1) the 1981 National Hospital Discharge Survey (NHDS) and (2) The American Hospital Association's (AHA) 1981 annual survey of hospitals. Data tapes for these two surveys were acquired for the year 1981 and population estimates for these samples compared to population estimates derived from the SMI sample. Overall, the SMI sample estimated 28,760,221 discharges for adult patients discharged from general hospitals in 1981 for other than normal delivery. The AHA estimated 35,714,899 admissions for all patients (a figure which may include normal deliveries, pediatric patients and admissions to subacute units). Finally, the NHDS estimated 34,532,257 admissions for specialty and general hospitals in 1981 for adult patients, a figure which may include patients admitted for normal deliveries.

The distribution of admissions by region was essentially identical for all three surveys. Approximately 28 percent of the admissions in each sample were in the North Central region, 20 percent in the Northeast, 35 percent in the South and 16 percent in the West. The average length of stay (ALOS) by region was also similar, with an average of six days in the West, seven in the South, eight in the North Central region, and nine in the Northeast.

The three populations were also comparable in terms of their distribution of admissions and days of care by bedsize and ownership. In addition, the distribution of this study's sample closely approximated that of the NHDS on a variety of patient characteristics including pay source, principal diagnosis, length of stay, patient age, patient sex, month of admission, month of discharge, the performance of major surgery, discharge status, and source of admission.

6. Estimation Procedures

The initial weight for a sampled record was the product of five weights, which took into account the probability of sampling the hospital, the probability of sampling a record within the hospital and three adjustments for nonresponse. In addition to the weight applied to the records, a weight was applied to the individual days of care.

The goal of the survey was to estimate the quantity $R = \frac{x}{y}$ where x equals the number of nonacute admissions or days, and y equals the total number of admissions or days. Sampling precision can be increased by using a ratio estimation procedure. Since x and y are apt to be correlated, this procedure increases the reliability of the estimate of x (nonacute admissions or days) by increasing the precision with which y (total admissions or days) is estimated. In the present study more precise estimates of total admissions and days of care were available from the National Discharge Survey, which sampled 226,585 patient records from over 400 hospitals in 1981. The NHDS data were used in the calculation of the first and second stage ratio estimators for admissions and days using a method adopted by the current population survey and described in Hansen, Hurwitz and Madow (1953). Since the NHDS includes specialty as well as general acute care hospitals, adjustments for the admissions and days of care contributed by specialty hospitals were made using the 1981 AHA Annual Survey of Hospitals.

7. Estimation of Sampling Variances

Given the complex multistage cluster sampling approach used to select hospitals, no direct computational procedure for estimating sampling variance could be applied. Therefore, a replication method of estimating the sampling variance of a statistic from multiple subsamples within the parent sample was used. This method was employed to obtain direct estimates of sampling variance for a number of quantities estimated in the analysis (e.g., the proportion of nonacute admissions in the nation as a whole, in the South, for Medicaid patients, etc.). These values were subsequently used to develop two generalization functions (one for admissions and one for days) which would allow sampling error to be approximated for any statistic calculated in the course of subsequent analysis.

8. Assessment of Inter-Rater Reliability

Since the SMI was extensively revised following the field test, an additional assessment of reliability was made. All records in seven randomly selected survey hospitals were reabstracted. An abstractor returned to the hospital and re-evaluated the same set of patient records that had been reviewed previously. The abstractor was one who had participated in the study, but was not the one who originally reviewed the records of the hospital. In all, 124 admissions and 1,245 days of care were reabstracted. The value of Kappa, an agreement measure which corrects the percentage agreement observed in the sample for the percentage that would be expected by chance, was .76 for admissions and .80 for days. Both these figures are significantly different from zero.

C. STUDY RESULTS

After analyzing sample representativeness and the reliability of reviewer evaluations, the study then estimated rates of nonacute hospital utilization and assessed factors believed to be associated with such utilization. These results are presented below.

1. The Magnitude of Nonacute Hospital Utilization

In 1981 a total of 1,925,000, or 5.7% of all admissions to acute care hospitals, were estimated to be nonacute. A total of 30,066,000 days, or 12.0% of all days, were found to be nonacute. Of the days, 4.2% (or 10,467,000) were unnecessary because the admission itself was unnecessary and 7.8% (or 19,599,000) were unnecessary days that occurred in otherwise justified admissions.

The rate of nonacute utilization found here is significantly lower than figures that have been reported in the literature. These figures have been as high as 18% to 20% for admissions and as high as 30% for days of care. Differences are thought to be associated with the limitations of previous methodologies and the inability to generalize from earlier studies based on small nonrepresentative samples.

The findings of this study -- that 5.7% of admissions and 12.0% of hospital days are unnecessary -- indicate that there is not as alarming a degree of hospital misutilization as has often been suspected. Although this study did not disclose blatant inefficiency in the use of hospital resources, the magnitude of unnecessary hospital use found here should not be viewed as trivial. If 12% of hospital days of care could be eliminated and/or redirected to lower cost settings, the potential cost savings are immense. Some inefficiencies in the delivery of health care are inevitable, of course, and cannot be controlled without incurring even greater economic and societal costs. Nevertheless, the findings of the study -- to be discussed below -- indicate that the majority of unnecessary care could be eliminated through changes in the behavior of hospitals, physicians, and patients.

2. Factors Associated with Nonacute Utilization

Analysis was performed separately for admissions and days of care. Only days falling in appropriate admissions were included in the day-level analysis. Within each category both descriptive and multivariate analyses were performed. The analyses were aimed at examining issues that had been raised in the literature on the subject of nonacute or unnecessary hospital utilization. These were as follows:

- what kinds of patient stays are most susceptible to nonacute use
- how much nonacute utilization is associated with inefficient scheduling or poor discharge planning

- how much nonacute utilization is associated with differences in hospital structure, resource availability, delivery style, management, and financial incentives
- how much nonacute utilization is associated with reimbursement mechanisms (especially insurance coverage) which encourage overuse of hospital care
- how much nonacute utilization is due to the unavailability of subacute care
- how much nonacute utilization is associated with area factors which might create induced demand
- how much of the regional variation in utilization rates (i.e., admission rates and ALOS) can be attributed to nonacute utilization.

One research question conspicuously absent is the degree to which nonacute utilization is related to variation in physician characteristics. The sensitivity of this issue precluded the study from obtaining any information on physician characteristics from participating hospitals.

a. Major Findings of the Admission-level Analysis

i. Descriptive analyses

Univariate analyses showed no significant differences in appropriate admissions by payment source (Medicare, Medicaid, Blue Cross, and other commercial insurance), or patient's discharge status. Although terminal cancer patients displayed a substantially higher rate of inappropriate admission than other patients, sampling error was large due to the small numbers of sampled patients in this category. The difference was thus not significant.

Admissions with characteristics indicative of more serious or life-threatening problems were more likely to be appropriate. These included patients placed in a critical care unit upon admission, as well as patients admitted through the emergency room versus those admitted routinely (after excluding those admitted for elective surgery). Older patients were also less likely to be inappropriately admitted, presumably because of the more serious nature of their illnesses. Similarly, patients with long lengths of stay (reflecting more serious problems) tended to have inappropriate admissions less frequently.

Inappropriate admissions conversely tended to occur in the non-vital organ systems, with the smallest portion falling in the class of circulatory disorders. The four diagnostic classes with a noticeably large proportion of inappropriate admissions, on the other hand, were (1) digestive problems, (2) musculoskeletal problems, (3) symptoms and ill-defined conditions, and (4) mental disorders. In each of these categories, the medical problems characteristic of nonacute admissions tended to be symptoms or signs of undetermined etiology (e.g., migraine headache, back pain, gastroenteritis, syncope) which can be incapacitating and/or painful and can be indicative of serious or life-threatening conditions, but which are not immediately life-threatening.

There was no tendency for patients hospitalized inappropriately to be discharged quickly (i.e., within one or two days), although they did have shorter than average length of stay. The highest percentage of nonacute admissions occurred for stays of from three to five days. These stays, in addition, accounted for 45% of all inappropriate admissions.

The highest percentage of inappropriate admissions were for diagnostic workup. In this category, 12% of the patients were admitted unnecessarily, compared with 7% for patients admitted in a medical treatment phase. Patients inappropriately admitted in a workup phase accounted for 41% to 47% of all inappropriate admissions whereas patients admitted for medical treatment that could take place in a subacute or outpatient setting accounted for 33% to 38% of all inappropriate admissions. Nonacute admission rates were lowest for those admitted in a pretreatment phase (i.e.; with a procedure, treatment or management regimen which has been determined but has not yet been initiated) and for those admitted into a surgical treatment phase. Only 2.5% and 1.6% of these admissions were inappropriate respectively. The latter, of course, does not indicate the appropriateness or medical necessity of the surgery itself, but merely the necessity of performing the surgery on an inpatient basis. Both, however, indicate that admissions for predetermined and definitive acute-level treatment are most apt to be appropriate.

A number of hospital and area level characteristics were associated with nonacute admissions. Nonacute admissions were significantly higher for smaller hospitals, hospitals with lower occupancy rates and hospitals with higher expenditures per bed. At the area level, nonacute admissions were significantly higher in rural areas, in areas with fewer physicians per capita, fewer nursing home beds per capita, and higher poverty levels. Additionally, nonacute admissions were significantly higher in the Southern region where admissions per capita are also highest and significantly lower in the Northeast where admissions per capita are also lowest.

ii. Multivariate analyses

Two types of analyses were performed in the multivariate examination of factors associated with nonacute utilization. The first analysis was at the patient level and the second at the hospital level.

Patient-Level Analysis. In the patient model, patient and stay variables, hospital variables, and area variables were examined. These were chosen both on theoretical grounds and on the basis of descriptive results. Since the appropriateness of an admission is a binary value, a logistic regression procedure was used to fit the model.

The effects that were significant in the logit patient-level model were similar to those apparent at the descriptive level. The patient and stay factors most associated with nonacute admissions were (1) the four diagnostic categories: digestive problems, mental disorders, musculoskeletal problems, and symptoms and ill-defined conditions; (2) placement in a medical, surgical or other non-critical care unit, (3) routine admission rather than admission through an emergency room (for other than elective surgery), (4) patients of younger age, and (5) the percentage of the stay devoted to diagnostic workup. Unlike the descriptive analyses, Medicare pay source was significant and associated with a lower proportion of nonacute admissions (an association which was accounted for by the significant effect of patient age). The dummy variable representing stays of three to five days was not significant owing to the fact that a somewhat different contrast was used in the logit analysis than the descriptive analysis.

In the patient-level logit model, patient and stay characteristics were the largest effects in the model. Patient and stay variables alone produced a rank correlation of .48 between the predicted probability and the actual value. The rank correlation for hospital characteristics alone was .16, and that for area characteristics alone was .22. The model based on the combination of patient, hospital, and area characteristics had a rank correlation of .55 between predicted and actual values. Each model was statistically significant.

Although there were a number of hospital and area variables which were significant at a bivariate level in the logit analysis, only a few loaded in the model due to the substantial collinearity among them. In the logit model constructed for hospital characteristics alone, expenditures per bed was significantly and negatively related to the probability of a nonacute admission. This effect overlapped the potentially significant effects of bedsize and occupancy rate. In the logit model constructed for area characteristics alone, hospitals in the Southern region and rural areas were more likely to have nonacute admissions and hospitals in areas with more hospital beds per capita were less likely to have nonacute admissions. In the final patient model logit containing all patient, hospital, and area variables, two regional effects were significant, namely, those representing the Northeastern region and the Western region, both of which had lower probabilities of a nonacute admission. The number of hospital beds per capita was also negatively associated with the probability of a nonacute admission.

Hospital level model. To perform the hospital-level analysis, the percentage of nonacute admissions was calculated for each hospital. Hospital and area characteristics were then regressed on this value. Hospitals were weighted to represent their respective strata. Eighteen hospital and area characteristics were identified as being the most theoretically-relevant and conceptually distinct. These were hospital bedsize, hospital expenditures per bed (indicative of resource intensity and resource availability), teaching status, proprietary ownership, occupancy rate, an ownership by occupancy rate interaction term, nursing home beds per capita, nursing home beds per aged population, physicians per capita, hospital beds per capita, the percentage of population below the poverty level, the percentage of physicians older than 55 years of age, rural location (defined as location in a non SMSA), and three regional dummy variables representing the Southern, Western, and Northeastern regions.

Nine variables were significantly related to nonacute admissions at a bivariate level. The percentage of nonacute admissions was higher for nonteaching hospitals, hospitals with fewer beds, hospitals with lower expenditures per bed, hospitals located in rural areas, in areas with fewer physicians per capita, in areas with higher poverty levels, and in the South. The percentage of nonacute admissions was lower in the West and Northeast.

Again there was substantial collinearity among the variables. That is, rural areas, which had smaller hospitals with lower expenditures per bed and lower occupancy rates, also had fewer physicians and hospital beds per capita, higher poverty levels, more older physicians, an older population, and were more often located in the South. This suggests that differences in nonacute utilization might be associated with a single regional or rural/urban dimension which reflects either differences in practice style between physicians in different geographic locations or differences in the availability of lower level facilities. Alternately, it suggests that regional and other geographic differences might be attributed to differences in the types of hospitals predominating in the area. Two forced order regression models were, therefore, estimated. The first order assumed that regional and geographic effects are the end product of the demographic characteristics and hospital-level processes which vary among geographic areas. The second order assumed that regional effects represent global differences in physician practices and patient expectations and that what occurs at the hospital level is the end product of these effects.

The hospital-level model was significant and accounted for 22% of the variance in the percentage of nonacute admissions. Only the effects of region, bedsize, older physicians and prevailing charges were significant when controlling for all other variables in the model. Although rural location was significant after controlling for region, it was not significant after controlling for hospital-level differences. Similarly, nursing home beds per capita contributed to the equation after controlling for some hospital and area variables, but did not have an effect when rural location and regional variables were entered in the equation. Occupancy rate was significant after controlling for region and other area variables, but not after controlling for hospital size and expenditures per bed. Teaching status was also not significant after controlling for hospital size and expenditures per bed. The effect of expenditures per bed, also appeared to overlap the regional effects substantially. The effects of older physicians and the prevailing charge rate, on the other hand, appeared only after controlling for all or a substantial number of other variables.

In sum, region had an effect on nonacute admissions above and beyond the effects of any of the demographic variables which vary between regions, suggesting that differences in regional style may account for the observed differences in nonacute utilization. In addition, hospital characteristics also made an independent contribution in explaining nonacute admissions. Of the various hospital characteristics examined, only bedsize made a contribution above and beyond those of other hospital and area characteristics. The causal role of occupancy rate, teaching status, and resource intensity are, therefore, unresolved. Finally, some variables, such as the prevailing charge rate and the percentage of older physicians in an area, are significant only after controlling for a number of countervailing tendencies, indicating that a number of forces may be operating in different directions within a given area.

b. Major Findings of the Day-level Analysis

i. Descriptive results

Medicaid patients had significantly fewer nonacute days. There were no significant differences found for the other major pay sources (i.e., Medicare, Blue Cross, and other commercial insurance). The patients most likely to have nonacute days tended to fall in the same diagnostic categories that were associated with nonacute admissions: namely, musculoskeletal problems, digestive problems, mental disorders, and symptoms and ill-defined conditions. This was true even though the day-level analysis included only days for which the admission was appropriate in the first place. Thus, even patients appropriately admitted with symptoms, back problems and so forth were, likely to incur a greater proportion of nonacute days. Additionally, in the day-level analysis patients with traumas (most likely nonpenetrating) were apt to have more nonacute days.

Not surprisingly, the types of stays most likely to have inappropriate days were longer stays, with the percentage of nonacute days increasing steadily as the length of stay increased from one to seven days. For stays over seven days, the proportion of nonacute days leveled off, but the absolute numbers of nonacute days in the stay continued to increase.

Patients discharged to skilled nursing facilities had significantly more nonacute days. The end of the stay also had the highest proportion of nonacute days (17%) and accounted for 43% of all nonacute days. The workup/pretreatment phase had the next highest proportion of nonacute days (11%), but accounted for only 16% of all nonacute days. Only 7% of the days occurring in a medical treatment phase were nonacute but these represented 40% of all nonacute days. The percentage of nonacute days in the active surgical treatment portion of a stay was negligible, accounting for only 1% of all nonacute days.

Surgical patients, (defined as any patient having a Class 1 procedure performed during the stay) tended to have proportionately fewer nonacute days. It is not clear whether this was true of the end of the stay, but in the pretreatment and the active treatment phase, surgical cases had a definite advantage. Of all nonacute days, only 9% were for surgical patients in a treatment, pretreatment or workup phase. Of these, 5% were patients in a diagnostic workup phase and 3% were patients in the pretreatment phase, i.e., those with a scheduled procedure pending.

Only 12% of all nonacute days were attributable to service delays (e.g., delays in ordering tests, scheduling surgery, receiving test results or consultations, etc.) which occurred when the patient otherwise required continued hospitalization, a category which included patients awaiting scheduled surgery. The highest percentage of days which were inappropriate for service-related reasons took place in a workup/pretreatment phase, but service delays still accounted for only 32% of all the nonacute days occurring in this phase. Of these nearly half (or 2% of all nonacute days) were attributable to service delays prior to surgery.

Of the nonacute days which took place, not because of service delays, but because the patient did not actually require hospitalization, 7% were for patients in a workup/pretreatment phase who could have had diagnostic tests performed on an outpatient basis. Twenty-four percent were attributed by reviewers to the unavailability of alternate care, and another 9% were due to such factors as expected improvements in patient condition that made discharge to a subacute facility unnecessary, delays pending patient education, and so forth.

The remaining 48% of days appeared to be potentially controllable. Of these, 15% were specifically attributed to physician or patient insistence by reviewers and another 10% to delays in discharge orders. There were, finally, 23% of nonacute days in the medical treatment phase or at the end of the stay for which the patient did not require institutional care, and reviewers were unable to assign a reason for the occurrence of the nonacute days. The patients incurring these days were not significantly different demographically from those who incurred nonacute days for which the patient/physician was assigned responsibility. That is, there was no evidence that these were overwhelmingly rural patients, elderly patients, or poor patients. Therefore, there is a strong indication that these unnecessary days are controllable and that they occurred through some combination of hospital, physician, or patient-specific actions.

Unlike the admission analysis, there were few hospital and area characteristics that were correlated with nonacute days. Among hospital characteristics, occupancy rate again demonstrated an inverse relationship, with the highest percentage of nonacute days for hospitals with the lowest occupancy rates. Region was significantly related to nonacute days, and the percentage of nonacute days by region corresponded to the observed regional variation in average length of stay: namely, the highest percentage of nonacute days was found in the Northeast where ALOS is the longest and the lowest percentage was found in the West where ALOS is the shortest.

ii. Multivariate analysis

As in the admission analysis, the strongest effects to emerge in the patient-level analysis were patient and stay characteristics. The logit patient-level model for patient characteristics produced a rank correlation between predicted and actual values of .50. The patient variables significant in the logit analysis were consistent with those found in the descriptive analysis. Namely, days at the end of the stay, days in a workup phase, days for patients discharged to a skilled nursing facility and days for patients with longer lengths of stay were more likely to be nonacute. Patients receiving surgery and those receiving critical care were less likely to have nonacute days (the former effect emerging strongly in the logit analysis but not in the descriptive analysis). As in the admission patient model, the effects of diagnostic category (digestive problems, mental disorders, musculoskeletal problems, traumas, and symptoms) were significant in various models. The effects were not as consistent or as stable, however, as those found for diagnostic category in the admission model.

In the day-level analysis, hospital and area characteristics had far less of an association with nonacute days than they do with nonacute admissions. The hospital-level model was not significant at all. In the patient-level logit model the rank order correlation of predicted probabilities with actual values was .12 for hospital characteristics and .07 for area characteristics. The rank correlation of the complete logit model containing patient, hospital, and area characteristics was .55, compared with .50 for patient characteristics alone. Additionally, in the logit model, the hospital effects were unstable. Occupancy rate was significant in the stepwise regression submodel for hospital characteristics but the effect was in the opposite direction from that in the descriptive analysis. Bedsize and expenditures per bed (which are highly correlated) variously emerged in different models. The regional effects, however, were stable and in the expected direction.

The small and/or unstable effects achieved by hospital and area characteristics are potentially a sign that physician and patient characteristics play more of a role in influencing nonacute utilization than hospital or area factors. Alternately the direction of hospital and area effects may differ depending upon the phase of the stay. Future analyses, therefore, might profitably attempt to specify different models for the different components of the stay.

D. DISCUSSION

Data analysis, described above, sheds some light on the factors associated with nonacute utilization. These are described below:

1. How much nonacute utilization is associated with inefficient scheduling and delays in discharge.

Only 12% of all nonacute days were attributable to delays in the provision of service. In addition, very few nonacute days appeared to be caused by the inefficient scheduling of surgery. Only 3% of all nonacute days occurred when scheduled surgery was pending and of these only 2% were actually due to delays in service. Additionally, the percentage of nonacute days that could be attributed to Friday admissions for Monday surgery was negligible. Only six-tenths of a percent of the patients in the sample were admitted on Friday and had major surgery performed on Monday. If all these patients (subtracting out a portion for which surgery was probably undelayable) had two unnecessary preoperative days, this would still amount to only 1% of all unnecessary days.

Nonacute days arising from delays in discharge may, on the other hand, contribute substantially to the problem of nonacute utilization. Forty percent of all nonacute days occurred at the end of the stay, and 27% were not attributable to placement problems, changes in patient condition, or other extenuating circumstances. There was, however, no evidence that patients held over the weekend for Monday discharge were more apt to have nonacute days (although due to small sample size, sampling error for the last day of the stay was large.) There was, in fact, a statistically higher than average percentage of discharges taking place on Friday and a statistically lower than average percentage of discharges taking place on Monday, suggesting that patients are more apt to be discharged on Friday than held over for Monday discharge.

2. How much nonacute utilization is associated with reimbursement mechanisms (especially insurance coverage) which encourage overutilization of hospital care.

Little information was available in this study that would allow detailed examination of this issue. The sampled numbers of self-paying patients and those covered by an HMO were too small, for example, to ascertain whether nonacute utilization was lower in these groups, both of whom had been hypothesized to have incentives for utilization control. In addition, the attempt to sample HMO-owned hospitals was unsuccessful. For patients covered by commercial insurance, there was no information as to the details of the benefit structure which would allow one to examine the effect of different types of incentives. It did not appear, however, that the Medicare and Medicaid programs were subject to any particular misuse with respect to hospital utilization. There was, in fact, no difference among the major payors (i.e., Medicare, Medicaid, Blue Cross, and other commercial insurance) in the proportion of nonacute admissions. There was, in addition, no difference among Medicare, Blue Cross, and other commercial insurance in the number of nonacute days. Medicaid patients, however, had a significantly lower proportion of nonacute days. This effect was not due to the younger age of Medicaid patients since age itself was uncorrelated with the proportion of nonacute days. This relationship was the only indication in the present study that the practices of various payors might influence the utilization of services.

3. How much nonacute utilization is associated with differences in hospital structure and organization.

In the admission analysis, a number of hospital characteristics were associated with nonacute admissions: in particular, smaller hospitals, nonteaching hospitals, those with lower expenditures per bed, and those with lower occupancy rates all had higher rates of nonacute admissions. The relationship between occupancy rate and unnecessary hospital admissions is particularly suggestive that hospitals themselves may play a role in inducing inappropriate admissions to compensate for low occupancy rates. Occupancy rate, however, is correlated with teaching status, bedsize and expenditures per bed and appears to have no effect above and beyond them. It is, therefore, difficult to draw any conclusions about the causal role of occupancy rate per se. Similarly, teaching status made no contribution beyond size and expenditures per bed.

These hospital characteristics were also correlated with a number of area characteristics including region, and urban/rural location. Few variables contributed after controlling for all others. Those that did, however, included both hospital characteristics (bedsize) and area characteristics (region). It appears, then, that hospital characteristics have some effect on nonacute utilization above and beyond area effects and, conversely, that the area influence on nonacute utilization is not solely mediated by the types of hospitals which predominate in an area. A causal modelling approach might be applied to better understand these inter-relationships.

In the day analysis -- where one would expect hospital characteristics such as resource availability, financial incentives, and so forth, to have the most effect -- no hospital characteristics were consistently related to nonacute days. This may, on the one hand, indicate that hospitals are less responsible for nonacute utilization than physicians (whose characteristics were unmeasured in the analysis). On the other hand, hospital (and area) characteristics may play a strong role in inhibiting or contributing to nonacute utilization, but this role may vary depending upon the phase of the stay. There is clearly potential for more finely-tuned analysis in the future, using less aggregated data.

4. How much nonacute utilization is associated with area characteristics indicative of resource availability, economic incentives, or practice style.

The hospital characteristics associated with nonacute admissions were also correlated with the area characteristics associated with nonacute admissions, indicating that the problems may not be hospital-induced but are endemic to an area. The characteristics associated with a higher rate of nonacute admissions were location in the South, in a rural area, and in areas with fewer physicians per capita, higher poverty levels, older physicians (in some analyses), fewer nursing home beds per capita (in some analyses), and fewer hospital beds per capita (in some analyses). These variables were all moderately to highly intercorrelated. This suggests possible effects of practice style differences, resource differences, or differences in economic incentives between areas.

a. How much nonacute utilization is due to the unavailability of alternate care.

According to reviewers' judgments, up to 18% of the nonacute admissions and 24% of the nonacute days were caused by the unavailability of subacute or home care. Thus, a significant portion -- though not the majority -- of nonacute utilization is attributable to the lack of alternate sources of care. There was little evidence to suggest, however, that nonacute utilization was affected by the supply of nursing home beds. In the admission model, nursing home beds per capita, although significant at a descriptive level, did not make an independent contribution to the model above and beyond all other variables. Similarly, there was no association at all between nursing home beds per capita and nonacute days.

Additionally, the higher rate of nonacute admissions in rural areas did not appear to be associated with the unavailability of alternate sources of care. The supply of nursing home beds per capita, for example, was not correlated with rural location. A scarcity of outpatient facilities similarly did not appear to account for the nonacute utilization occurring in rural areas.

b. How much nonacute utilization can be attributed to supplier-induced demand.

Although there was a strong suggestion in the data that the majority of unnecessary admissions are controllable by physicians and their patients, there was only the most indirect evidence, of supplier-induced demand -- either of physician supply or the supply of hospital beds. Nonacute days were not associated with supply characteristics at all. Nonacute admissions were, in fact, negatively associated, being lower in areas with more physicians and hospital beds. McClure (1982) hypothesizes that the high utilization rates that occur in areas where physicians and hospitals are concentrated reflects -- not inappropriate utilization -- but utilization which lies within the range of acceptable practice. That is, the surplus admissions contributing to higher admission rates are not medically inappropriate, but merely reflect more "elaborate" utilization philosophies which are nevertheless professionally defensible. Alternately, higher utilization rates may merely reflect higher rates for elective surgery, the necessity of which would not be measured by the Standardized Medreview Instrument.

Nonacute admissions, then, varied inversely, rather than positively with physician and hospital supply variables, indicating that the factors promoting nonacute utilization are greater in low-supply areas than they are in high-supply areas. There was, however, one finding in the admission analysis which suggested that when these influences are controlled, physician-induced demand may play a role in nonacute admissions. This was the positive relationship between nonacute admissions and the prevailing charge rate for twelve major procedures. This effect was significant only when controlling for all other variables in the model, and indicated that, all other things being equal, nonacute admissions will be higher in areas where physicians are paid more for a hospitalization.

In addition, although nonacute utilization was negatively associated with physician and hospital bed supply and had no association with per capita utilization rates at a county level, it did appear to explain much of the regional differences in utilization rates.

c. How much nonacute utilization is associated with differences in practice style.

The observed differences between rural/urban areas and regions in the proportion of nonacute admissions may also in part or as a whole be attributable to differences in patient expectations and physician practice styles. In the hospital-level model, the rate of nonacute admissions was lower in areas with a greater percentage of older physicians, after controlling for hospital-specific factors, health resource supply factors, and other area factors. Very few physician characteristics were available for the analysis, and none was available at the patient-level. Although this relationship is quite possibly mediated by some unmeasured factor, it is one of the few pieces of evidence to suggest that physician practices and characteristics themselves may be related to differences in nonacute utilization.

The strongest evidence that differences in nonacute utilization are caused by differences in local practice style, however, are the regional differences in nonacute admissions and days which exist even after controlling for a number of other factors which vary among regions.

d. How much of the regional variation in utilization rates can be attributed to nonacute utilization.

Aside from patient and stay characteristics, regional differences were the strongest effects found in the study both for the admission and day level analyses. The pattern of nonacute utilization found by region was consistent with observed variation in admission rates and average length of stay. Namely, the rate of nonacute admissions was highest in the South where admissions per capita are also highest and lowest in the Northeast where admissions per capita are also lowest. Similarly, the percentage of nonacute days was highest in the Northeast where average length of stay is highest and lowest in the West where average length of stay is also the lowest.

Regional effects were significant in explaining nonacute admissions even after controlling for a variety of hospital and area characteristics which were correlated with region. Regional effects were significant in explaining nonacute days even though there were no other hospital or area characteristics associated with nonacute days. This would seem to argue, first, that observed regional variation in admission rates and average length of stay are partially attributable to overutilization and, second, that such tendencies are at least partly due to differences in practice style or philosophy, rather than to demographic and resource differences that vary across regions.

E. CONCLUSIONS AND IMPLICATIONS

The Standardized Medreview Instrument is based on criteria established and in widespread use for PSRO review at the time the study was initiated. In the course of instrument development the criteria were enhanced to incorporate revisions to the original criteria set that had been made by over 50 PSROs. These revisions were reviewed and integrated by a physician panel. The criteria were further streamlined and clarified based upon the results of a fifteen-hospital field test. In the survey, reviewers achieved more than acceptable levels of reliability in applying the criteria. In addition, the reviewer evaluations were extensively investigated following data collection and were found to have been made in accordance with the intention of the criteria set.

An independent evaluation of the Standardized Medreview Instrument and its properties relative to both other utilization review methods and established medical judgment is a high priority in the continued effort to understand the phenomenon of nonacute hospital utilization. Based on the lengthy development and evaluation process undergone by this instrument, however, we feel, (1) that the criteria of the Standardized Medreview Instrument are in accordance with accepted standards of medical practice and utilization control and (2) that the criteria were applied by reviewers in a manner consistent with their intent.

In this national study, 5.7% of all admissions were found to be nonacute. Of these nonacute admissions:

- Up to 18% were patients who might have been inappropriately hospitalized because of the lack of availability of alternate care (including patients with no home care and terminal patients hospitalized for humanitarian reasons).
- Another 6% were patients who needed subacute care, but were hospitalized at patient, family, or physician insistence.
- A total of 76% patients required no institutional care at all. Of these only 36% were definitely attributed by the reviewer to patient or physician insistence. For another 40% no reason could be deduced. However, these patients were not demographically different from those hospitalized at patient or physician insistence.

With respect to days of care, this study found that 12.0% of all days and 8.1% of all days occurring in admissions that were justified were nonacute. Of the latter:

- Up to 24% were attributable to the unavailability of other forms of care.
- 9% were due to changes in patient condition, the need for patient education, and anticipated improvements in patient condition.

The remaining 67% of nonacute days appeared to be potentially controllable:

- 12% were due to delays in service, tests, consultations and so forth, of which 2% took place while a patient with a scheduled procedure was awaiting surgery. Whether all 10% of these days could be eliminated or whether some represent inevitable inefficiencies occurring in the medical decision process is unclear.
- 7% were for diagnostic testing that could have been performed on an outpatient basis.
- 25% occurred at the end of the stay or in the medical treatment phase and were definitely attributed to physician/patient insistence or to delays in discharge order.
- 23% occurred at the end of the stay or in the medical treatment phase for patients who required no subacute care at all. Reviewers were unable to assign responsibility for the nonacute days but these nevertheless appear to be under the control of the patient, physician or hospital in that these patients had no alternate care needs and were not demographically different from patients retained at physician or patient insistence.

In sum, this study found a lower rate of inappropriate utilization than has generally been accepted. While 5.7% inappropriate admissions and 12.0% inappropriate days is lower than the figures currently reported in the literature, it is still a nontrivial amount in terms of costs to both federally- and privately-funded health care programs. Using the estimates of this study, 30,066,000 days of care for adults hospitalized in general acute care hospitals in 1981 could potentially have been eliminated and the care provided in other than an acute care setting.

Not all these days can, of course, be easily eliminated. Nevertheless, allowing for (1) the 18% of nonacute admissions which are due to the unavailability of alternate care and the 24% of nonacute days in justified admissions which occur for the same reason, (2) the 9% of nonacute days which are due to patient condition and/or patient education factors and even (3) another 10% nonacute days which occur through inherent inefficiencies in the medical decision process, the data still suggest that well over half the nonacute days could be eliminated without major structural changes to the existing health care system.

With approximately one-fifth of all nonacute days attributable to the lack of alternate care, there is also substantial potential for cost-savings to be made by providing the mechanisms and facilities for alternate care. In addition, the 5.7% unnecessary admissions and 12.0% unnecessary days of care found in this study should not be taken to imply that there are no cost savings to be gained beyond this level. First, the study in no way touched upon the cost savings that might be possible through more efficient use of ancillary services in the hospital. Second, the necessity of elective surgeries was not assessed in this study. Nineteen percent of the admissions in the study were for major (i.e., Class 1) surgery which were defined as elective in that the patient did not otherwise meet criterion signifying the need for immediate intervention. If even 5% of these are unnecessary, the number of unnecessary admissions is increased by almost one percent.

Finally, and more centrally, the criteria of the Standardized Medreview Instrument might be considered as defining the upper boundary of what McClure (1982) terms the "range of acceptable practice." Within this range, as McClure points out, there are physicians, such as those in many prepaid group practices who make "conservative use of inpatient physician and hospital services" while others make aggressive use of these services. The magnitude of unnecessary hospital use found in this study, thus, does not reflect the reductions possible from practicing a more "conservative" style of medicine or those to be gained from widespread implementation of newly-developing alternatives to hospital-level care. The 5.7% inappropriate admission rate, for example, does not include the 1.7% of procedures that were performed in the hospital, but which could have been performed in an ambulatory surgery unit if one were available.

The percentage of hospital admissions and days over which we have control is thus bound to decrease as society searches for less costly alternatives to acute-level care and as these innovative approaches to alternate care become standard practice.

F. SUGGESTIONS FOR FURTHER RESEARCH

With approximately one-fifth of nonacute days occurring because of the unavailability of alternate care, it appears there is a substantial potential for cost-savings to be made through providing the mechanisms and facilities for alternate care. The findings so far do not shed much light on the kinds of resource deficiencies which contribute to this problem. No definitive relationship was found between nonacute use and nursing home beds per capita at the county level. It is not clear, therefore, whether the problem is due to the sheer unavailability of beds or is simply one of coordinating patients with available beds. Other issues, such as whether a subacute unit attached to the hospital promotes efficiency, also remain unexplored.

Another set of questions remaining unanswered relates to the causal mechanisms through which nonacute utilization occurs. The present study did not provide any definitive answers. However, it is suggested that more fine-tuned analysis of the day-level data set along with a causal modeling analytical approach might still make a contribution in this area. In this endeavor it is not the intention of the study to place blame on any one factor of the health care system. However, knowing the circumstances under which nonacute utilization occurs is crucial in understanding what actions must be taken in order to control it.

In one sense, the answer to this question is irrelevant. With the introduction of the prospective payment system under Medicare and the possibility of widespread adoption of this system, society is coming closer to placing the burden of controlling nonacute utilization (at least for days of care) on the provider. Nevertheless, health care utilization is a complex phenomenon and the prospective payment system and other regulatory and competitive strategies may have a number of unintended consequences. Efforts to gain a better understanding of the processes which contribute to nonacute utilization can only help payors and providers alike to work together to provide high quality, cost-effective care.

Suggestions for future research include first, an independent evaluation of the Standardized Medreview Instrument, comparing it with other utilization review instruments as well as independent physician judgment. In addition, further analyses of this data set should be conducted at a more refined level, including (1) more detailed analyses of the day level data set to identify factors associated with nonacute days by phase of stay, (2) a causal modeling analysis of factors associated with both nonacute admissions and days and (3) an examination of nonacute utilization arising specifically from the unavailability of alternate care.

In addition, if similar data collection efforts are conducted in the future, collecting more records per hospital and more detailed information on hospital characteristics would be worth considering. Similarly, collecting data on physician characteristics would considerably enhance the explanatory potential of the analysis. A real understanding of the processes and causal factors involved in nonacute utilization may, however, only be attainable through studies done on-site concurrently with the patient stay.

II. INTRODUCTION AND BACKGROUND

II. INTRODUCTION AND BACKGROUND

The escalating costs of health care are well documented. These have risen from 4% of the Gross National Product in 1950 to almost 10% in 1981 (U.S. Bureau of the Census, 1982). Hospital costs have risen even faster than other types of health care costs. As hospitals have increased in number and sophistication, they have accounted for an ever-larger portion of national health expenditures. In 1929, 18% of the dollars spent on health care were spent in hospitals; by 1978 this proportion had increased to 45.2% (Levit, 1983). The amount currently spent for hospital care is over twice that spent on physician services, the next largest category.

Although increases in utilization have played a significant role, inflation in hospital prices has made the major contribution to this situation. Whereas the Consumer Price Index (CPI) rose 288% between 1965 and 1981 and that for medical care as a whole rose 326% in this period, hospital room charges increased by over a factor of six (U.S. Bureau of the Census, 1982).

Improved wages for hospital personnel and general inflationary pressures have been responsible for some of the increase in hospital costs. Other factors, however, are also involved. The widespread adoption of new and more costly technologies -- while possibly contributing to shorter length of stay -- is thought to have increased costs overall (Mechanic, 1978). A compounding factor is the increased demand for health care -- in particular, more technologically sophisticated levels of health care -- brought about partly by rising incomes but more so by the almost universal expansion of insurance coverage. In this regard, direct patient payments for all medical care declined from 52% in 1965 to 32% in 1981. Further, for hospital care taken alone, third party payments accounted for 89% of total expenditures, (U.S. Bureau of the Census, 1982).

Because patients are insulated from considerations of price in determining their health care needs, patients, physicians, and hospitals have had little incentive to forego more expensive forms of health care even when lower cost alternatives are available and would result in the same benefits to health status. Insurance plan coverage and exclusions, in fact, often provide negative incentives for using the most cost-effective treatment modality. This has been compounded by the fact that until recently, nearly all hospital revenues were billed on a per diem or per service basis. Thus, hospitals have had few financial incentives to limit services.

All these factors have operated to maintain a health care system which is not only high cost and technologically intensive, but which would appear to discourage the efficient use of these resources. Many of the proposals to contain health care costs have, therefore, focused on identifying and controlling what has been variously termed unnecessary hospital utilization, inappropriate utilization, overutilization or misutilization. In this effort, researchers and policy-makers alike have been concerned with assessing the degree to which unnecessary hospital utilization exists and the factors that are responsible.

A. THE MAGNITUDE OF UNNECESSARY HOSPITAL UTILIZATION

1. Assessment of Inappropriate Utilization Through Comparison of Utilization Rates

That some portion of hospital use is unneeded has long been suggested by studies demonstrating significant differences in hospital utilization rates across regional and other geographic boundaries, as well as across various practice settings. Regional (Detmer and Tyson, 1978) and national (Bunker, 1970) variations in surgery rates have been noted, as well as differences in all other aspects of utilization. For example, hospital days of care per 1000 Medicare enrollees have been found to vary by as much as 36% between regions (Gornick, 1976). A more recent study (Gornick, 1982) has confirmed these differences, showing significant variation in both the discharge rate and average length of stay across regions. In 1977, days of care varied 43% across regions (being highest in the North Central region and lowest in the West); the discharge rate varied by 22% across regions, and average length of stay varied by 49%. Regional variation of these two components of utilization is largely uncorrelated: For example, the discharge rate was highest in the South and North Central region (354 per 1000) and lowest in the Northeast (301 per 1000) and West (317 per 1000). Average length of stay was highest in the Northeast (13.3 days) and North Central region (11.1 days) and lowest in the West (8.9 days) and South (10.0 days). These trends remained after controlling for age, sex, race, and diagnosis. The same pattern was also found for patients less than 65 years of age. Other studies (Blumberg, 1982) have shown that the regional effect remains after standardizing by reported morbidity rates. McClure (1982), in fact, reports that controls for differences in demographic characteristics and health status generally account for no more than 15% of the variance in utilization.

Variations in hospital utilization have been studied using small-area analyses which examine utilization among ostensibly comparable populations living in neighboring localities. These have found wide variations in the incidence of surgery (Roos, Roos, and Henteleff, 1977; Wennberg and Gittelsohn, 1973, 1975, 1982; Stockwell and Vayda, 1979; Lewis, 1969) and hospital utilization (Wennberg and Gittelsohn, 1973, 1982; Wennberg, Gittelsohn and Soule, 1975; McCracken, 1980). McCracken (1980), for example, examined both admission and days of care rates for Iowa counties, and found utilization rates ranging from 30% below to more than 50% above the statewide mean, after adjusting for age-sex composition. In examining 193 areas in six New England states, Wennberg and Gittlesohn (1982) found the rate of variation of surgical procedures to vary by a factor of two, after adjusting for age-sex differences. For specific procedures such as hysterectomy or tonsillectomy there was variation by up to a factor of six. Within state variations in hospital discharge rates ranging from 127 to 235 discharges per 1000 have been found in Maine (Wennberg, Gittelsohn, and Soule, 1975) and from 122 to 197 discharges per 1000 in Vermont (Wennberg and Gittelsohn 1982). All these studies took into account place of patient origin to avoid cross-area biases.

In the attempt to identify population characteristics which might account for these differences, Wennberg and Fowler (1977) interviewed residents of six areas in which surgical rates varied by more than a factor of two. They found little difference in morbidity, the proportion of people living under the poverty level, the proportion of people with health insurance, the proportion with access to a physician or the proportion visiting a physician each year. They concluded: "The large differences in surgical rates and amount spent on hospital care must, therefore, be traced to factors that come into play after patients have contact with physicians." (Wennberg and Gittlesohn, 1982).

Organizational differences have also been shown to affect hospital utilization rates. Admission rates for prepaid group practices are as much as 25% to 45% lower than fee-for-service admission rates, and individual practice associations have been shown to have rates from 0% to 25% lower (Luft, 1978; Roemer and Shonick, 1973). In at least one instance, fee-for-service group practices have also been able to achieve hospital utilization rates equivalent to group HMOs. Nobrega, et al., (1982) studied the Mayo Clinic and the Olmsted Medical and Surgical Group, finding 1976 age-sex adjusted hospital discharge rates to be 38% less than the national rate. The study attributed these differences in rates of hospital utilization to the types of physician practice patterns accepted and promoted within a particular delivery system, and concluded that low utilization can be maintained in a fee-for-service setting while providing high quality care.

The possibility has been raised that the lower utilization rates found in HMOs reflect not efficiency but rather underutilization due to lower quality of care. This quality of care supposition has generally not been supported by research, regardless of whether quality has been measured along structural, process or outcome dimensions (Luft, 1980). In addition, studies such as that by Nobrega, et al. tend to confirm that high quality care is not inconsistent with lower rates of hospitalization.

Thus, these variations in utilization rates across settings are not explained by demographic differences, differences in health status or in the quality of care. Such variation has, however, been closely tied to the supply of health resources, notably both hospital beds and physicians per capita. Bunker (1970) noted that the number of surgeons per capita was twice as high in the U.S. as it was in England, as was the rate of surgical procedures. Vayda and Anderson (1968) showed a similar relationship for Canadian provinces, Lewis (1969) for counties in Kansas, and Wennberg, et al., (1982) for 193 hospital areas in New England. Roemer and Shain (1959) first introduced the concept of supplier-induced demand, namely, that an increase in the supply of hospital beds will increase per capita utilization, not through the price mechanism, but through a supplier-induced shift in the demand function. Other studies (e.g., Anderson, 1973; Harris, 1975) have not only found a positive relationship between bed supply and utilization but have found that an increase in per capita bed supply is positively related to subsequent utilization. These studies are subject to a number of methodological difficulties that limit the interpretation of causality (see Wilensky and Rossiter, 1983; Ginsburg and Koretz, 1983). However, recent analyses which control for many of these methodological problems have still shown a positive effect attributable to supply-induced demand (Ginsburg and Koretz, 1983).

In sum, variations in utilization rates that are tied more strongly to supply factors than factors indicative of medical need suggest that a significant portion of hospital utilization may be unnecessary, and, therefore, inappropriate. It is apparent from these statistics that a wide range in discretionary utilization exists. However, it is not clear that discretionery utilization necessarily implies unnecessary utilization or at what point one turns into the other. McClure (1982), for example, hypothesizes that utilization rates from as low as 600 to 800 days per 1000 (that achieved in the most efficient HMOs) to as high as 1400 to 1600 days per 1000 both lie within the range of "acceptable practice" in that they produce equal outcomes and adhere to standards professionally acceptable to the medical community. Attempting to identify and define unnecessary hospital utilization solely through a comparison of utilization rates, therefore, has a number of drawbacks. This has led to a number of attempts to measure unnecessary use directly at the patient level.

2. Direct Assessments of Inappropriate Utilization

More direct evidence confirms the existence of some level of unnecessary hospital utilization, although estimates of the exact percentage of unnecessary care vary considerably. One of the most frequently cited areas of concern is elective surgery. Some second opinion surgical screening programs have demonstrated a significant level of unnecessary surgical utilization for elective procedures. In New York, McCarthy and Widmer (1974) found one-fifth of Storeworkers Health and Welfare Fund participants, and one-third of District Council members did not have a recommended surgery confirmed in a second opinion surgery program. Later studies found similar results (Ruchlin, et al., 1982; Grafe, et al., 1978; McCarthy and Finkel, 1978; Gertman, et al., 1980). The Colorado Foundation for Medical Care, on the other hand, evaluated the necessity of thirteen common surgical procedures performed in the state in the latter half of 1978 and found only 3% did not meet explicit indications or were not justified by physician review. Others have also questioned the existence of such high rates of unnecessary surgery (e.g., Sammons, 1977).

The assessment of other types of hospital use has been less clear cut. In a national survey of administrators conducted in 1976-77, Rothberg and Gertman (1981) found that only 2.1% of hospital administrators and 5.5% of Utilization Review coordinators felt that there was no unnecessary utilization, while 6.8% and 12.7% respectively thought over 15% of all days were inappropriate. Overall, the mean percentage of inappropriate days was estimated to be 5%. Studies of hospitalizations using direct evaluative methods have, however, found a wide range in levels of inappropriate use, particularly with respect to bed-days. Gertman and Bucher (1969), for example, reported 11.9% of the acute medical days at the Baltimore City Hospitals to be unnecessary according to five explicit guidelines developed for the study. Restuccia and Holloway (1976), using the Medicare levels of care criteria to define appropriate utilization, found 10.6% of all inpatient days at a medium-sized California community hospital to be inappropriate. Zimmer and Groomes (1969) found only 2.0% of the bed-days to be inappropriate using an approach that required three observers to agree in their evaluation, but found 18.2% to be inappropriate if any unfavorable evaluation classified that bed-day as inappropriate. Zimmer (1967) similarly found that the estimate of inappropriate bed-days ranged from 1.7% when four observers were required to agree, to 26.6% for any unfavorable evaluation. Berg, et al., (1969) found that observer evaluations of unnecessary bed-days varied from 14.3% when local physicians made these judgments under the assumption that existing rather than ideal alternate health facilities were available, to 23.8% when visiting physicians made these judgments and assumed that optimal health facilities existed. Finally, even higher levels of inappropriate utilization were found using Gertman and Restuccia's (1981) Appropriateness Evaluation Protocol (AEP): two nurses and one physician each reviewed 200 charts at Boston University Medical Center Hospital and concurred that 30% of the days of care for these cases were inappropriate.

While these studies as a whole imply that some level of unnecessary hospital utilization exists, several limitations are apparent. First, none was carried out on a representative sample. Thus, the results cannot be generalized to the national, regional, or even local level. Second, each study employed different methods of defining and measuring unnecessary use. In addition, as the studies by Zimmer (1967), Zimmer and Groomes (1969), and Berg, et al., (1969) show, the levels of inappropriate use found in a study are very much influenced by the methodology used. Finally, as will be discussed further in Chapter III, these evaluative methods were generally subject to high levels of unreliability. Thus, as noted by Kurylo (1976), the rates of misutilization observed by these studies cannot be generalized with any degree of confidence. As she notes, "...the inescapable conclusion is that misutilization measurement requires far more research before it can validly and reliably be used to promote better resource use."

B. CAUSES OF UNNECESSARY HOSPITAL USE

A key element in the elimination of unnecessary hospital use is the identification of the causes of unnecessary hospital use and an assessment of the degree to which these factors are controllable. Empirical studies of utilization behavior unfailingly demonstrate that utilization is affected by many variables not related to the need for health care, such as insurance coverage, resource availability, patient demographics, and so forth. (McKinlay, 1972; Veeder, 1975; Ro, 1969; Harris, 1975; Rosenthal, 1965; Brewer and Freedman, 1982; Lave and Leinhardt, 1976). While often suggestive, these models have not been specifically directed toward identifying the factors contributing to unnecessary utilization (see, however, Health Policy Institute, 1980). In addition, as in all studies which examine aggregate utilization, it is difficult to distinguish among underutilization, discretionary utilization, and overutilization.

The evidence from small area studies and studies of utilization in different practice settings implies that much overutilization is attributable to supply-side factors that are controllable by physicians and/or hospitals. However the mechanism through which these effects operate is not readily apparent. Different organizational structures, practice settings, and incentive systems have been suggested, but no definitive association has been found. Additionally, although unnecessary utilization takes place with the acquiescence or insistence of patients, it is not clear whether there is any way to predict which patients will tend to overutilize, or to identify characteristics (e.g., insurance coverage) responsible for this tendency. Other factors beyond the control of hospitals, physicians, or patients, most particularly the availability of alternate facilities for care, may also be associated with unnecessary utilization.

The following discussion presents the existing evidence with respect to a variety of factors potentially implicated in unnecessary utilization as well as presenting evidence concerning regulatory efforts to control unnecessary utilization.

1. The Physician's Role in Inappropriate Utilization

Since physicians are the major decision-makers in any patient stay, many researchers have hypothesized that they are the agents most responsible for inappropriate utilization, (Pauley and Redisch, 1973). Although evidence is limited, there is some suggestion that this is the case. Restuccia and Holloway

(1976) in a study of the "Barriers" to appropriate utilization at a California community hospital found that 41.5% of all inappropriate days were considered by reviewers to be physician caused. This included such factors as physician delay in ordering tests or physician delay in ordering discharge.

Other studies have found consistent differences among physicians in utilization rates even after controlling for casemix. Horn, et al., (1980), for example, examined 29 staff physicians at a medium-sized community hospital. They identified 11 physicians who had longer than average length of stay, controlling for diagnosis, and five physicians with average or less than average length of stay. When patient records for these two groups of physicians were reviewed by the hospital's utilization review department, they were found to differ significantly in the average number of unjustified days per patient stay.

Borchardt (1981) has similarly found that most of the inappropriate utilization in the cases reviewed by the Delmarva Foundation for Medical Care was attributable to physician, rather than hospital or patient factors. In addition, a small number of physicians can account for the majority of inappropriate days at an institution. Using the Appropriateness Evaluation Protocol (AEP) to calculate a nonacute profile for 239 physicians, Borchardt found the percentage of inappropriate days to range from 0% for some physicians to over 50% for others. Feedback of these findings to physicians succeeded in reducing the percentage of nonacute days. Although physician variability in the tendency to overutilize is apparent, it is not clear what characteristics distinguish overutilizers from other physicians. Borchardt noted a difference by specialty with physicians in Thoracic surgery, Neurology, Plastic surgery and Psychiatry at the top of the list in terms of the percentage of inappropriate utilization. Horn, et al. found no differences between the physicians identified as high utilizers and the comparison group in terms of physician age, length of practice, or attendance at meetings.

The literature on prepaid group practices discussed previously suggests that the organizational setting of care can greatly impact the utilization of hospital services. Whether this is due to the different economic incentives operating in a fee-for-service environment as opposed to a group practice environment (Eisenberg and Rosoff, 1978) or to other organizational or practice style differences is still debated (McClure, 1982). Scitovsky and McCall (1980), for example, argue that the peer pressure of a group practice may serve to control unnecessary utilization even in a fee-for-service practice. Nobrega, et al., (1982) have, in fact, observed lower than average utilization in two fee-for-service group practices. Noting that the fee-for-service reimbursement system of these groups offers no incentives to limit utilization, they attribute this to the "motivation, practice philosophy, and organization of the two clinics themselves."

The observation that utilization is higher where the rate of physicians per capita is also higher has long suggested that physicians may instigate unnecessary hospital use in highly competitive areas in order to maintain income levels. More recent evidence suggests that this may be a more complex phenomenon than previously thought. Wilensky and Rossiter (1983), used interview data collected at the patient level, rather than area-wide aggregate data to study this issue. After controlling for patient insurance coverage, they found some evidence that physicians induced ambulatory follow-up visits, but none suggesting any physician inducement of elective surgery. McClure (1982) suggests that the observed relationship between physicians per capita and utilization rates may be due to self-selection, such that physicians with naturally "elaborate" utilization styles are able to compete in high supply areas and, therefore, remain in these areas, whereas physicians with "conservative" utilization styles are unable to compete and so gravitate to more rural areas or to prepaid group practices.

2. The Effects of Insurance Coverage

Insurance coverage has always been a powerful predictor of utilization behavior. Numerous studies demonstrate that the insured use more services than do the uninsured, and utilization varies directly with the amount of coverage held (Newhouse, 1978a). In the past, these differences have been attributed to lack of access and underutilization on the part of the uninsured. More recently, concern has focused on the possible negative incentives inherent in the present insurance system which may encourage the unnecessary and inefficient use of health care (Newhouse, 1978b, Heaney and Reidel, 1970).

This issue has, for example, been raised in relation to the observed rate of increase in hospitalizations among the aged since the implementation of Medicare coverage in 1965. As noted by Lubitz and Deacon (1967) the rate of discharges for persons over 65 years of age rose by 35% between 1967 and 1979 whereas that for persons under the age of 65 rose only 11%. Noting the concurrent rise in physicians and bed supply per capita, these researchers raise the possibility that the phenomenon may be due to unnecessary utilization deliberately induced by physicians to maintain their income in an increasingly competitive environment.

The problem may be more complex, however. Lubitz and Deacon note that the hospitalization rate also increased by 22% in the same period for persons 45 to 64 years old. This suggests that "technical advances, perhaps spurred by the advent of Medicare coverage, may have been adopted at a more rapid rate for conditions among the elderly than for other conditions." Additionally, while the admission rate rose for Medicare-covered patients, average length of stay declined over the same period, so that total days of care per 1,000 eligibles remained roughly constant.

Gornick (1982) notes that the major fear at the time the Medicare and Medicaid legislation was enacted was that utilization would increase substantially, either through increased access or through abuses of the programs by providers. Little concern was expressed that price or unit costs might increase. As she states: "It is near-legendary now that it was costs -- rather than patient demand as measured by the number of physician office visits or hospital days of care -- that spiraled the first year these programs were implemented and for several years to come." Much of this cost inflation has been attributed to the cost-based reimbursement provisions of the Medicare program, an attribution which provided the impetus for the recently implemented Medicare prospective payment system.

The form of insurance coverage and the method of reimbursement, therefore, appears to have powerful, if subtle, effects on the delivery style and pricing behavior of providers. Consumers have additionally been insulated from considerations of the costs associated with their own health care decisions or those made in their behalf by health care providers. There has consequently been increasing interest in instituting cost-sharing arrangements with consumers as a means of reducing unnecessary utilization. (Heaney and Reidel, 1970). Proposals have been made in two areas: the first is directed toward increasing the cost-sharing provisions for hospital care and the second is directed toward decreasing the cost sharing provisions for ambulatory care (which historically have been much less comprehensively covered than inpatient care) in the hope of encouraging the substitution of ambulatory for inpatient care.

A number of studies of different cost-sharing provisions have been made. These, summarized by Chassin (1978), indicate that hospital utilization is insensitive to cost-sharing arrangements and is reduced only with the application of unacceptably high cost-sharing rates. Ambulatory care, on the other hand, is quite sensitive to cost-sharing, with increased ambulatory coverage leading to increased ambulatory utilization. The evidence is not clear, however, as to whether this increased ambulatory care substitutes for hospital care, except in an HMO setting.

Many of these studies are also methodologically flawed by the fact that assignment to a specific plan was not random, yielding a high probability that low utilizers will be self-selected into plans with high cost-sharing provisions and conversely for high utilizers. A recent study conducted at Rand (Newhouse, et al., 1982) used an experimental design, with random assignment of families to plans that varied in terms of the coinsurance rate, i.e., the fraction of medical bills the family was required to pay. The same coinsurance rate applied to both inpatient and ambulatory services except for one plan in which the coinsurance rate applied only to ambulatory services with inpatient services being free. Maximum out of pocket expenditures were limited to $1,000 per family. Preliminary results are consistent with those summarized by Chassin. In particular, the study found that as cost sharing decreases, the percentage of persons seeking care, the number of ambulatory visits per person and the number of adults hospitalized increases. When cost-sharing was applied only to the ambulatory setting, with all inpatient care free, hospitalization rates actually declined. These investigators also concluded that cost-sharing does not affect costs per hospital stay particularly given a ceiling on out-of-pocket expenditures of $1,000 per family as was the case in that experiment. Given that the findings were not particularly promising with respect to the substitutability of ambulatory for hospital care, it may be necessary to design insurance coverage so as to explicitly exclude inpatient coverage for care that should be obtained on an outpatient basis.

3. Hospital Characteristics Related to Inappropriate Utilization

The literature on the hospital characteristics associated with costs and utilization is far-ranging, comprising numerous theoretical discussions and empirical investigations of the organizational factors affecting hospital costs and efficiency (e.g., Feldstein, 1967; Berki, 1972; Lave and Lave, 1970; Lave, Lave, and Silverman, 1972; Ro, 1969; Goldfarb, Hornbrook, and Rafferty, 1980). Included among the organizational characteristics studied are many with potential significance for the analysis of inappropriate utilization.

Numerous factors have been hypothesized to affect hospital efficiency. Among others, hospital size (Feldstein, 1967), profit ownership (Rushing, 1974), and membership in a multi-hospital system (Coyne, 1982; Zuckerman, 1979) have all been thought to promote efficiency, either through economies of scale or economic incentives. Teaching hospitals, on the other hand, have been hypothesized to be less efficient and cost conscious. Various internal administrative and organizational structures have also been studied for their effects on efficiency (Longest, 1974; Scott, et al., 1979). The evidence in all areas has been inconclusive.

These investigations additionally provided only limited insight into the hospital characteristics that might be related to unnecessary utilization. Efficiency, as defined in such studies, may in fact translate into the more efficient and timely delivery of services, thereby decreasing utilization. However, cost efficiency is not necessarily related to control overutilization, particularly since financial incentives may operate in the reverse direction. Occupancy rate -- or excess bed capacity -- has been suggested to be a key variable in stimulating unnecessary utilization (Health Policy Institute, 1980). In hospitals with low occupancy rates, the organizational characteristics promoting cost consciousness (e.g., profit ownership) might, in fact, produce higher rates of inappropriate utilization, as hospital administrators seek to maintain revenues.

Studies have been performed relating hospital characteristics to average length of stay. These are suggestive, but difficult to interpret. As noted by Lave and Leinhardt (1976), casemix differences account for the largest component of the variance between hospitals in ALOS. After applying controls for casemix, residual differences in ALOS theoretically represent differences in unnecessary utilization. It is difficult, however, to be certain that all casemix differences have been accounted for by whatever adjustment procedure is used. Ro (1969), for example, found, after adjusting ALOS for differences in diagnosis that the number of facilities in the hospital was positively related to residual ALOS, as was the presence of an intern or residency program. The number of employees per bed was negatively related to residual ALOS, as was the presence of a home care department and the presence of an outpatient clinic. No effect was found for occupancy rate. These findings, thus, lend support to the wisdom that teaching hospitals are more inefficient while hospitals which have mechanisms for providing alternate care (i.e., home care department and outpatient clinic) are more efficient. They additionally suggest some efficiencies associated with having a larger staff to patient ratio. Nevertheless, the positive correlation of adjusted ALOS with the number of services might indicate that the casemix control was incomplete: the longer ALOS of more service intense hospitals might, therefore, reflect a more severely-ill patient population.

Given the inherent ambiguity of aggregate utilization measures, this issue could profitably be studied using direct measures of inappropriate utilization. In one of the few studies performed on-site concurrent with the patient stay, Restuccia and Holloway (1976) found little inappropriate utilization to be due to hospital factors. Of all inappropriate days, 42% were considered to be physician-caused and 42% due to environmental factors such as the unavailability of alternate care. Only 7% were attributed to delays in the provision of service or inadequate discharge planning on the part of the hospital. Zimmer (1974) similarly found only 8.1% of inappropriate days to be due to a delay in a hospital procedure. (No attribution as to physician or hospital responsibility was made, however. Additionally, the hospital role in inadequate discharge planning was not assessed).

Similar studies are required on a larger sample of hospitals before conclusions can be drawn as to the magnitude of the hospital's contribution to unnecessary utilization.

4. The Availability and Appropriate Utilization of Alternate Care Facilities and Services

It is generally recognized that a substantial portion of hospital misutilization occurs for patients who require alternate sources of care, either at the institutional level (e.g., long-term care hospital, SNF, or other nursing home) or through home care programs. This problem has been observed in this country, as well as in Britain (Hall and Bytheway, 1982; Rubin and Davies, 1975) and Canada (Shapiro and Roos, 1981). It is particularly thought to affect geriatric patients, the terminally-ill, and the socially disadvantaged.

Studies of this problem have all been based on data from a single hospital or local group of hospitals. Therefore, little definitive information is available as to the magnitude of the problem, the characteristics of those affected, whether misutilization is due to the sheer unavailability of alternate care resources or whether it is due to the failure to use these resources in a timely and efficient manner. These studies, however, uniformly indicate that a nontrivial portion of hospital misutilization occurs for patients who should be receiving services at an alternate level.

a. The Magnitude of the Problem

In a probability sample of institutionalized patients over 65 years of age in Monroe County, New York, Wenkert, et al., (1969) found that 7% of these patients had been inappropriately placed in general acute care hospitals. Similarly, in a probability sample of bed-days covering all general hospitals in Monroe County (Berg, et al., 1969), observers judged that 18.8% of occupants did not require care or could have been cared for in other existing facilities. Of these, 40% were judged to require long-term hospital, nursing home, or organized home care.

Rosenfeld, et al., (1957), studying four Boston general acute care hospitals, found that 6% of all patients remained in the hospital 30 days or longer. These patients accounted for almost one-third of the total patient days rendered. On 42% of these days, the patient did not require continuing active treatment in the hospital, although many required other types of care. The main reason these patients were kept in the hospital was attributed to lack of alternate care, an explanation which encompassed both the unsuitability of the patient's home or lack of other facilities and services. Greenberg and Rogers (1961), who reviewed patients hospitalized for calendar year 1961 at New York Hospital, found that 6.5% of the patients were hospitalized for over 30 days. These patients accounted for 25% of the bed-days. A survey of attending physicians indicated that 28% did not need acute-level care. Had discharge taken place sooner, 10% of the bed-days used by these patients (or 2.5% of all bed-days) would have been eliminated. The lack of alternate care facilities appeared to be the major factor for the retention of these patients, with 61% of the patients ultimately discharged to another facility incurring inappropriate days. More recently Lynch (1980) found that in a one-month period, 784 patients in Connecticut acute care hospitals waited 6,034 hospital days for placement, for an average of 7.7 extra days per case. This figure represented 2.3% of all hospital days.

b. Characteristics of Patients and Patient Stays

Although the investigations discussed above have focused on long stay cases, the problem also occurs in a significant way for short stay cases. Zimmer (1974), for example, studied patients hospitalized at Strong Memorial Hospital of the University of Rochester and found that, overall, 9.1% of the bed-days were inappropriate. Observers concluded that on 74% of these days, the patient required alternate care. The alternate care problem was clearly more pervasive for longer stays: Only 9% of days in stays over thirty days were for patients not requiring any care at all. This figure was 44% for stays of one to seven days. Although of lower magnitude, it is clear that even for short stay cases, the unavailability of alternate care may be a significant problem.

The characteristics of patients "at risk" for alternate placement problems conform to expectations. Greenberg and Rogers (1961), for example, found that one-third of patients remaining hospitalized when discharge to a subacute care facility was appropriate were over 65 years of age. Additionally, one-third were receiving terminal or pre-terminal care and 14% died while in the hospital. The group at greatest risk were women over 65 years old, without a husband, with poor prognosis, and inadequate home conditions. Financial characteristics did not appear to play a role. Glass, et al., (1977) found that four characteristics predicted the likelihood of a patient's incurring "social" days (i.e., days on which the patient required lower level care, but was not discharged). These were (1) age greater than 80 years old, (2) the need to live somewhere new at the completion of the stay, (3) the occurrence of any form of disorientation, and (4) chronic disorientation. These factors were highly predictive of the number of social days incurred (r = .46). The ability to ambulate and to perform the activities of daily living did not explain any of the variance beyond these factors. Financial characteristics could not be measured for inclusion in this scale.

c. The Contribution of Planning Versus Resource Availability

The studies described indicate that much hospital misutilization occurs because of the difficulty of finding alternate care for the functionally and socially impaired population. In many hospitals, however, such as that studied by Glass, et al., (1977), there is an organized and active social work staff who plan the dispositions of all patients prior to discharge. In such hospitals one would, therefore, expect most inappropriate days to occur because of extreme placement difficulties, rather than inefficiencies in planning for discharge. In a different setting, Restuccia and Holloway (1976) found that only 41% of all inappropriate days and only half the days at the end of the stay were attributed to environmental conditions, such as the unavailability of SNF care, home care, or alternate care for terminal patients. Although this is a substantial figure, it indicates that some delays in discharge -- even for patients requiring subacute care -- may be avoidable.

Thus, while the difficulty of placing certain patients in alternate care facilities contributes substantially to the problem of hospital misutilization, it is not clear whether the problem is one of sheer unavailability of alternate care resources or one of coordination. In this light, Ingman and Claus (1981) argue for better preadmission assessments of geriatric patients and propose a method based on the problem-oriented medical record. Berkman and Rehr (1973) conducted an empirical assessment of early case finding using an experimental design. They found that among those patients who ultimately needed help in seeking placement, average length of stay was ten days shorter for patients brought to the attention of social workers upon admission.

Some studies have, in addition, found no noticeable effect of the level of alternate care resources in an area on inappropriate utilization. Gertman, et al., (1979) in evaluating the early impact of PSROs found no effect of either nursing home turnover rates or nursing home occupancy rates on Medicare days of care or length of stay. Shapiro and Roos (1981) similarly found that transfers of long-stay elderly patients to long-term care facilities in Manitoba took longer in 1976 than in 1972, despite major provincial initiatives which occurred in the interim and included the construction of additional long-term treatment beds, expansion of home care resources and extension of universal insurance coverage to long-term institutional and home care. These investigators concluded that the major problem was the lack of central planning which resulted in discretion over admissions by the facilities concerned. That is, long-term care facilities could delay admissions of less desirable patients or otherwise select patients to fit their financial and behavioral needs.

It is also clear that in designing the optimum health care system, a variety of forms of alternate care must be provided. In the study by Greenberg and Rogers (1967), for example, physicians indicated that much of the social stay problem was due not only to the unavailability of acceptable nursing homes, chronic care facilities, and special facilities for the terminally-ill, but that other services or facilities were needed. These included (1) hotel type accommodations close to the hospital for preoperative workup, special diagnostic studies and convalescence, and (2) an overnight ward for 24 to 48 hour stays to rule out acute emergencies such as appendicitis or myocardial infarctions as well as for special diagnostic procedures.

5. The Impact of Utilization Review Programs

During the past two decades numerous regulatory strategies have been implemented with cost containment as the goal. The success or failure of such strategies in controlling costs and/or utilization may be useful in indicating the factors involved in the inefficient use of resources. The literature on this subject is extensive and numerous studies, most of them inconclusive, have been conducted (see Chassin, 1978; Stuart and Stockton, 1973; NCHSR, 1977; Sloan and Steinwald, 1980). This section will discuss only the attempts to control utilization through the PSRO movement and related utilization review programs.

Utilization review (UR) evolved during the 1960s as a method of controlling hospital utilization, and therefore reducing costs. In addition, physicians and hospital administrators saw utilization review as a technique to ensure quality of care. It was first recommended (1963) and then required (1966) by the Joint Commission on Accreditation of Hospitals (JCAH) as a requirement for the certification of hospitals (Chassin, 1978). In the same year, utilization review became mandatory for hospitals receiving Medicare and Medicaid reimbursement. The requirement arose in part from legislative concern that the newly implemented Medicare and Medicaid programs would result in overutilization of health care by these newly enfranchised populations. By law, hospitals participating in either program were to review both "duration of stay and professional services furnished" (Federal Register, 1966). In 1972, PL92-603 incorporated these UR requirements into a more comprehensive program of inpatient review that established the Professional Standards Review Organizations (PSRO) program. Under this law, both control of hospital utilization and quality assurance monitoring were to be addressed by the PSRO utilization review process.

Specifically, PL92-603 identified three areas of PSRO responsibility for reviewing services delivered and paid for under Medicare, Medicaid, Maternal and Child Health, and Crippled Children's Programs. Services were to be: (1) medically necessary; (2) of a quality that met professional standards; and (3) provided at the most economical level consistent with quality of care (Goran, 1979). Using these criteria, the focus of PSROs was on concurrent continued stay certification and retrospective profile analysis. Concurrent admission review and prospective admission review, i.e., preadmission certification, were added to the program subsequently.

a. Evaluations of Utilization Review Programs

i. Local and area-wide evaluations

Positive effects of utilization review programs have been found for single institutions. For example, Westphal, et al., (1969), using accounting data over a five year period for a South Carolina University Hospital, showed a significant decrease in length of stay and average charges per patient for Medicare and Medicaid patients immediately following implementation of concurrent PSRO admission and length of stay review. No similar decrease was found for patients not covered by the review. Similar success in lowering ALOS has been found at a medium-sized community hospital in Massachusetts, even after controlling for casemix (Wymelenberg, 1978).

Some area-wide evaluations have also shown a decrease in admission rates and/or ALOS. Fulchiero, et al., (1978), for example, found a decrease in ALOS for Medicaid patients in Massachusetts in the two and one-half years of operation of the Commonwealth Hospital Admissions Monitoring Program. A comparison group of non-Medicaid patients did not exhibit the same downward trend even after adjusting for age, sex and diagnosis. Brook, et al., (1978), on the other hand, found no demonstrable impact of the New Mexico Foundation for Medical Care on utilization for the Medicaid population statewide. There was no significant downward trend in rates of hospital admission, length of stay, or total days of care over the four year study period. Disease-specific time series regression models also showed no effect attributable to the program.

Other studies which have shown an apparent effect of utilization review programs have been criticized for methodological weaknesses (Chassin, 1978; Davidson, Wacker and Klein, 1973; Sayetta, 1976). Brian (1973), for example, examined the impact of the Certified Hospital Admission Program (CHAP) in Sacramento county and the impact of the state utilization review effort in two other PSRO areas. These programs involved both preadmission and continued stay certification for the state's Medicaid patients. The analysis showed that CHAP reduced average monthly days of hospital care 16.7% below predicted levels, while the state effort achieved decreases of 13.5% and 9.5% in two other counties. Sayetta (1976), however, points out that statewide ALOS was already declining and the trend was well established before the UR program went into effect. No acceleration of the trend was apparent. The comparison population used by Brian showed a downward trend similar to that of the Medicaid population, and was not well-matched demographically in any respect.

Similarly, Flashner, et al., (1973) found that the Hospital Admission and Surveillance Program (HASP) in Illinois, a program which approved length of stay on a prospective basis, decreased ALOS by one-day and also decreased total Medicaid reimbursements to hospitals. Davidson, et al., (1973), however, note that Medicaid reimbursement procedures changed at the same time and the lowered reimbursement rates arising from this change explained the slowed rate of growth in Medicaid payments. These investigators also criticized the control data used by Flashner, et al. on a number of grounds.

In a subsequent evaluation of HASP, Wylie and Flashner (1979) used better control data, matching Medicaid patients to a comparison group on age, diagnosis, and severity of illness. Although ALOS decreased in this period, it fell significantly more for Medicaid than non-Medicaid patients, indicating a positive effect of the program. In addition, casemix changed more markedly (in the direction of increased severity) for Medicaid patients than for non-Medicaid patients.

In addition to other methodological problems noted above, many of these early evaluations were flawed in that they concentrated on average length of stay and cost per patient, without also considering the admission rate and total days of care. In principle, ALOS should decrease as inappropriate days are eliminated by utilization review. However, a successful admission review program will also eliminate the less severely ill, who have both shorter ALOS and lower costs per case. Given this possibility, average length of stay and average cost per patient may actually increase rather than decrease. Therefore, a more appropriate measure for evaluating the effect of specific UR programs is to examine rates of admission and total hospital days for the covered population, adjusting for changes in casemix and demographic characteristics (Chassin, 1978; Boaz, 1979).

ii. National evaluations

National assessments of the PSRO program have highlighted the variable impact of the program. The Health Service Administration's Office of Planning, Evaluation, and Legislation OPEL (HSA, 1977) examined changes in Medicare and Medicaid utilization between 1974 and 1976. They found no aggregate effect on utilization of PSRO areas versus non-PSRO areas, but did find that some individual PSROs were associated with favorable impacts, some with no impact, and some with a negative impact.

A comprehensive nationwide evaluation was performed by HCFA (HCFA, 1979, 1980) to examine the effect of PSRO program review over the period of 1974 to 1978 on Medicare discharges and ALOS. A national evaluation could not be conducted for Medicaid patients due to the lack of uniform data. This study found no significant impact of the PSRO program overall, but did find an impact which significantly varied by region. In the Northeast, PSRO review explained 3.5% of variance in days of care (primarily attributable to decreases in ALOS). In the North Central region, PSRO review explained 1.4% of the combined effect of ALOS and admissions. In both these areas the sign of the impact was favorable, i.e., PSRO review reduced days of care, after controlling for region. In the Northeast, PSRO review had a significant negative impact on ALOS, but not on the rate of admissions. In the North Central region, PSRO review reduced total days of care. Neither the admission rate nor ALOS were significantly affected. However, for both the effect of PSRO review was in the negative direction. In the West, PSROs had no impact on days of care, ALOS or discharges, and in the South, PSROs had an impact which was significant but in the opposite direction, i.e., in areas with active PSRO review, days of care were higher. In total, the PSRO program appeared to save 1.7% to 2.0% days of care in active PSRO areas. A number of caveats were expressed by the authors, including the fact that the program had not been fully implemented as of 1978, particularly in the South. Inactive areas in the South were highly dissimilar from active PSRO areas. Therefore, the results of the analysis did not necessarily imply that expansion of the PSRO program in the South would continue to produce seemingly unfavorable results.

Despite some methodological deficiencies which limit interpretation of these various studies, it appears that utilization review programs have a variable impact, which can be positive in some areas or for some hospitals, neutral in others, and even negative in some instances. The cause of these differences -- whether due to area-wide influences or mediated by the involved hospitals -- has not yet been established. Rothberg, et al., (1980) point to the variety of patient, demographic, and hospital factors which vary among regions and might explain utilization differences as well as the differential impact of the PSRO program; these investigators, however, offer no conclusions.

iii. Attributes of effective review programs

Gertman, et al., (1978) note that there were significant utilization review activities already in place in many hospitals at the time the PSRO program was initiated, a fact not taken into account by the OPEL evaluation. These investigators identified hospitals which had an operating UR program that had effected at least one admission review termination in 1976. Admission termination was chosen because admission review was at that time a voluntary activity not required for Medicare patients and so indicated that the hospital's program was actually committed to the function of utilization review. Such hospitals experienced a decline in Medicare days of care and ALOS from 1974 or 1976, whereas hospitals without a UR program or with a program that had not effected any admission terminations experienced an increase. There was no effect of a hospital's being under binding PSRO review or having a conditional PSRO operating in the area. Thus, committment to utilization review as signified by admission terminations was a more significant factor than the formal nature of utilization review.

Other investigations such as that performed by Restuccia (1982) have examined the organizational qualities that enhance the effectiveness of individual UR programs. He found a significant reduction in inappropriate days through concurrent review. The effectiveness of review, however, was influenced by the type of feedback the UR coordinator was allowed to use. In particular, concurrent review was more effective when UR coordinators were allowed discretion over when and with whom they communicated, and could speak with the attending physician directly without necessarily involving the physician advisor.

Other studies have focused on the temporal aspects of the review process which might enhance or impede the effectiveness of review. These have focused primarily on the effectiveness of concurrent continued stay review, concurrent admission review and preadmission certification. Studies by investigators such as Rosenfeld, et al., (1957); Greenberg and Rogers (1967); and Zimmer (1967) described earlier, showed that a high proportion of inappropriate utilization (primarily attributable to placement problems) occurred for long stay patients in the latter part of the stay. This resulted in an emphasis on extended duration stays and certification/recertification procedures (Goldberg and Holloway, 1975), using LOS percentiles to indicate the point at which review should be performed.

Although longer stays are more likely to have inappropriate days and inappropriate days occur more frequently at the end of the stay, studies have shown a substantial amount of misutilization occurs in short stays and in other parts of the stay. For example, Zimmer (1974) found the previously observed effect of length of stay: Namely, for those hospitalized for one to seven days at the time of review, 5% of the days were judged to be inappropriate, whereas 25% of the days were inappropriate for patients hospitalized over 60 days. In absolute numbers, however, Zimmer found that short stay cases accounted for the majority of inappropriate days.

Restuccia and Holloway (1976) similarly found that the percentage of inappropriate days increased with the number of days in the hospital, from 2.3% on the first day to over 15% on days fifteen, sixteen, and seventeen. Again, however, short stay cases accounted for the greatest proportion of inappropriate days: almost 40% of inappropriate days occurred in stays of less than eight days. Regardless of length of stay, the end of the stay had the most nonacute days: 65% of the inappropriate days occurred at the end of the stay, 10% in the initial stage, and the remaining 25% in the middle part of the stay. A substantial number of inappropriate days, however, occurred earlier in the stay. Twenty-one percent occurred before the fiftieth percentile for a given diagnostic category and 37% occurred before the 75th percentile.

As a consequence, Averill and McMahon (1977), among others, question the cost-effectiveness of continued stay certification, given the inevitable time delays involved in reviewing the patient, making a decision and then instituting remedial action. These researchers developed a model to predict the number of patients who would be reviewed under current timing guidelines, and would also experience a sufficiently large number of inappropriate days for the review to effect a reduction in costs. Using a variety of assumptions, they found dubious evidence that a sufficient number of inappropriate days could be eliminated through concurrent review to justify the cost of the program. As a result, they argued that the most efficient use of resources involved retrospective profile analysis to identify problem areas on which subsequent concurrent review efforts could profitably be focused.

In a similar vein, Boaz (1979) argues that preadmission certification is a more effective approach to utilization containment than concurrent or continued stay review, citing evidence from regression models that the effects of nonmedical factors on length of stay are small and concentrated among patients whose diagnoses might not qualify them for hospitalization under a preadmission screen in the first place. Boaz's definition of cases which might not qualify for admission, however, was liberal and included 60% of all patients, namely those not classifiable as urgent or emergent admissions.

Dittman and Magee (1980), recognizing that each form of review might be appropriate for different situations, developed a comprehensive model combining the three approaches in the optimum manner. As a final caveat applicable to all programs aimed at controlling utilization alone, Studnicki and Honemann (1982) warn that hospitals can vary utilization along two different dimensions: duration (i.e., LOS) and intensity (i.e., the delivery of diagnostic and therapeutic services). Control over duration will not necessarily lower the total costs of hospitalization without control over intensity, a consideration which argues in favor of the DRG-based prospective payment system over that of concurrent stay review.

C. SUMMARY

In summary, wide variations in hospital utilization rates (including discharges, surgeries, and average length of stay) have been found across various geographic and organizational settings. These differences cannot be attributed to demographic factors or differences in quality of care, but do appear to be related to differences in physician and bed supply. This strongly suggests that a non-trivial portion of hospital care may be eliminated without any decrement in the health status of the population. Given the high costs of hospital level care, efforts to identify and control unnecessary utilization should be undertaken.

Estimates of the percentage of hospital care appropriately provided at a lower treatment level are difficult to obtain. Equally difficult to ascertain is the portion of such care which is avoidable through changes in the behavior of hospitals, physicians, and/or patients, rather than structural changes in the health care system. Conclusions based on aggregate utilization statistics are ambiguous at best. Direct measurements of inappropriate utilization made at the patient level are indicated, but these have not produced consistent results. The lack of a standard, uniformly accepted methodology for measuring the levels and causes of inappropriate utilization has hampered theoretical and practical efforts to address these issues. Kurylo (1976), in reviewing the direct evaluative measurements of inappropriate utilization available at the time the project was conceptualized, found them to be untested and largely unreliable. In addition, such measures as had been used were applied to small and/or nonrepresentative samples.

This study was designed to correct these methodological deficiencies. It is the first assessment of inappropriate hospital utilization using a national probability sample. In addition, the initial part of the study was devoted to the development and rigorous assessment of a reliable instrument to measure inappropriate hospital utilization. These two phases, namely the development and testing of a method to measure unnecessary hospital utilization and the application of that method in a nationally representative sample, are described next.

In subsequent discussions, the term "nonacute" has been used in the place of "inappropriate" (following a convention introduced by Borchardt, 1981). This was to avoid the pejorative connotations of the term "inappropriate", particularly since some portion of this utilization is undoubtedly outside the control of hospitals or physicians -- and may be appropriate on social grounds -- even if not justified by the medical need for acute care services. Since the goal of the study was to identify all instances where acute care hospitals are being used for other than the delivery of nonacute level services (regardless of the social or medical appropriateness of the care) "nonacute" was felt to be the more descriptive term.

III. DESIGN OF THE STUDY

III. DESIGN OF THE STUDY

This project was initiated in July 1980 as part of a program of research by the Health Care Financing Administration/Department of Health and Human Services. The study, completed in September 1983, had two components:

- the development of a methodology for measuring nonacute hospital utilization through the synthesis of existing approaches, and
- the application of the method on a nationwide scale to estimate the magnitude of nonacute utilization and identify the factors associated with it.

In implementing each of these phases, the project was guided by a Technical Advisory Panel (listed in Appendix A) which identified sources of information, aided in the conceptual and operational definition of nonacute utilization, and provided advice on instrument evaluation and on the design and analysis of the national survey. In addition to informal consultation, the panel met periodically to review project activities. The instrument development and the national survey phases of the study are discussed in turn below.

A. DEVELOPMENT OF THE METHODOLOGY

1. Overview

The concept of "inappropriate utilization" implies that the care is either altogether unneeded or, more commonly, that the care could be delivered in another, less resource-intensive, setting. The philosophical approach taken by the study was to define the appropriateness or necessity of hospitalization in terms of medical necessity alone, i.e., by whether the medical circumstances of the patient's condition require the use of hospital-level resources, both upon admission and throughout the stay. Given this definition, if a patient could be treated in a lower-level setting or at home although requiring the help of family, friends, or other home health services, hospitalization is to be judged inappropriate (even if no lower-level care is available or the patient had no family willing or able to care for him). The same was considered to be true of any other hospital day which occurs -- not because it is medically necessary for the care of the patient -- but because of external constraints (such as the inability to arrange transportation or the inability to discharge the patient due to adverse environmental conditions). A similar argument relates to problems of resource availability in the hospital which might be beyond the hospital's physical or financial control, (e.g., rescheduling of surgery due to an emergency or delays in treatment caused by the unavailability of weekend laboratory staffing).

This definition was not meant to imply that hospitalization in these instances is socially, as opposed to medically, unjustified, or that the patient should pro forma be discharged. The goal was instead to identify all instances in which hospitals are used, not for the provision of acute care, but to provide other services to society. In this way, the cost of various alternatives might rationally be weighed against the cost of the existing problem.

In addition to distinguishing socially necessary from medically appropriate utilization, the study also distinguished the concept of appropriate (meaning necessary or efficient) utilization from a number of other problems that may occur with respect to hospital care. For example, problems that may best be considered ones of under-utilization or quality of care (such as delayed admission or premature discharge) were not defined as examples of inappropriate use in the present investigation. (The methodology eventually developed for the study does, however, allow identification of those patients who require continued hospitalization but are not receiving services and/or are discharged prematurely.)

The present study in no way attempted to assess whether the specific acute-care services provided during a stay were medically indicated for the patient's condition or whether the patient's problem was managed correctly. Finally, the study was concerned only with appropriateness of the acute care setting itself, and not with assessing ancillary or other service utilization within the hospital (e.g., unnecessary provision of diagnostic tests or overuse of special units). Although methods for assessing ancillary service utilization were available at the time of project initiation (most notably the APACHE or Acute Physiologic and Chronic Health Evaluation system for evaluating the need for critical care, Knaus et al, 1981) incorporation of these methodologies was considered beyond the scope and objectives of the study.

In developing the methodology, three situations leading to unnecessary utilization were identified: namely, unnecessary utilization was hypothesized to occur (1) when a patient is hospitalized unnecessarily (as in an admission for a procedure which could be performed on an outpatient basis or for treatment which could be rendered in a subacute nursing facility or in a physician's office), (2) when the patient is kept in the hospital after he/she is ready for discharge or transfer to a lower level of care and (3) when services are delivered inefficiently, so that the hospital stay is unnecessarily prolonged (as when the patient is kept in the hospital several days before a procedure is performed). Accordingly, a requirement of the methodology was that it be able to evaluate the necessity of the admission separately from the appropriateness of continued stay, and if possible, to distinguish a continuation of stay that is unwarranted on the basis of the patient's condition from one that is prolonged through inefficient delivery of services.

With respect to the evaluation of days of care (as opposed to the admission itself) it was considered essential to be able to assess the necessity of each day individually, rather than to judge the appropriateness of the stay as a whole based on length-of-stay norms. The ambiguity inherent in the use of length of stay norms was discussed in the previous chapter: length of stay is confounded with differences in severity of illness and so inferences based on length of stay lead to misleading and sometimes contradictory conclusions. In addition, such an aggregate measure does not allow one to identify the part of the stay in which unnecessary days occur and so does not allow one to identify the causes of (and potential solutions to) nonacute hospital use.

At the time of project initiation, extensive work had been performed in the area of utilization review and in the development of methodologies for assessing the appropriateness of hospital utilization. Given this, it was this project's assignment to draw upon existing methods to the greatest extent possible in operationalizing the concepts outlined above. Accordingly, the first step in the process of methodological design was to examine the current status of utilization review methods. This was done by reviewing the literature and by contacting researchers and others concerned with utilization review methods in practical settings. The following section discusses the status of these methods at the time the study was initiated. This is followed by sections discussing (1) the preliminary instrument and the steps taken for its evaluation and refinement and (2) the final methodology developed for the national survey.

2. The Status of Utilization Review Methodologies at Project Initiation

Interest in utilization review and evaluation was first evident in the 1960s. It was recommended (1963) and then required (1966) by the Joint Commission on Accreditation of Hospitals (JCAH) for the certification of hospitals (Chassin, 1978). Whereas physicians and administrators viewed utilization review primarily as a component of quality assurance, legislative concern at this time focused on the use of utilization review to ensure that the newly implemented Medicare and Medicaid programs did not encourage the over utilization of health care services.

As a result, utilization review became mandatory for hospitals receiving Medicare and Medicaid reimbursement in 1966. Under this requirement both quality assurance and nonacute utilization were combined in a review of "duration of stay and professional services furnished" (Federal Register, 1966). In 1972, PL92-603 further integrated these two aspects of utilization review into a more comprehensive program of inpatient review by establishing the Professional Standards Review Organizations (PSROs), which were to ensure that services delivered and paid for under Medicare, Medicaid, the Maternal and Child Health and Crippled Children's Programs were: (1) medically necessary, (2) of a quality that met professional standards and (3) provided at the most economical level consistent with quality care (Goran, 1979). Using these criteria, utilization review performed by PSROs consisted of either concurrent review (admission certification and concurrent continued stays certification) or retrospective profile analysis. Quality assurance was generally provided by means of retrospective medical care evaluation studies (MCEs).

These legislative actions served as an impetus for increased interest in the process of utilization review. As a consequence, the specific methods for conducting utilization review and targeting unnecessary utilization have evolved greatly over the last two decades. During the late 60s and early 70s, length-of-stay criteria were the most frequently used indicator of inappropriate utilization (Brian, 1972; Flashner, 1973), an approach that was supported by early hypotheses that the greatest amount of unnecessary hospital utilization occurred at the end of the stay. Utilization review methods accordingly relied on length of stay guidelines to indicate the appropriateness of continued hospitalization and to indicate the timing of continued stay review. Early methods of Medicare/Medicaid review often used a single value for length of stay, using the "period of extended duration" as the point at which any hospitalization would be subject to review. This system was then refined on a diagnosis-specific basis. For concurrent review programs, PSRO guidelines recommended the use of the length of stay findings of the Professional Activity Study (PAS) conducted by the Commission on Professional and Hospital Activities (CPHA). These length of stay norms are based on U.S. census region averages and are assigned to patients based on admission diagnostic information. Continued stay certification (CSC) then examines patients who have exceeded a certain percentile, typically the 50th or the 75th, for such norms. The 50th percentile has frequently been used, as stays beyond this point were believed to be most likely to include unnecessary utilization (Burford and Averill, 1979).

The validity of length of stay as an indicator of appropriate use has been questioned on a number of grounds. Diagnostic specific norms are often not sensitive to levels of severity and, in fact, if admission review is successful, average length of stay may increase due to the elimination of the nonacutely ill. Zimmer and Groomes (1974) reviewed 3,369 patients of Strong Memorial

Hospital at the University of Rochester and concluded that, in absolute numbers, the greater portion of inappropriate utilization occurred in short-stay cases. Similarly, Restuccia and Holloway (1976) found the number of inappropriate days to be negatively correlated with length of stay. Further, a number of studies have indicated that a substantial amount of unnecessary hospital utilization occurs in the early part of the patient stay (Gertman and Bucher, 1971; Zimmer, 1974).

A further objection to the use of length of stay norms for utilization review derives from the fact that Professional Activity Study (PAS) length of stay norms are based on discharge data, whereas Continued Stay Certification reviews patients placed in PAS diagnostic categories based on admission data. Burford and Averill (1979) have shown norms based on admission diagnosis to be inaccurate in 41.3% of the cases examined in 68 hospitals. The study found differences between admission and discharge diagnosis to be a substantial obstacle to the application of PAS length of stay norms for utilization review purposes.

On the basis of such evidence, interest focused on more direct ways to measure inappropriate utilization. These efforts have attempted to define the conditions and services which distinguish the acute-care setting from other care settings. Goldberg and Holloway (1975), for example, define this solely in terms of the services provided:

> Each patient (and each diagnosis) may be thought of as possessing a "services profile" which changes during the total duration of a hospital stay. Some of these services can be provided only in a hospital, others may often be provided in a hospital (among other possible sites), and still others are not ordinarily provided in a hospital. The services profile changes as the number, intensity, and type of services shift over time. As the shift occurs, there will usually be fewer services that can be provided only in a hospital, for the patient discharged alive, until there are finally no services which absolutely require a hospital setting...By focusing on the first two categories of services (hospital only, and hospital among other possible sites), explicit criteria can indeed be generated for the type, number and intensity of services, irrespective of diagnosis, which constitute a hospital level of care.

Goldberg and Holloway (using a list prepared by the McKay-Dee Hospital Center in Ogden, Utah) propose a number of factors to be considered in establishing the need for hospital level services including both patient condition and service factors: these are severity of illness, stage of illness, presence of complicating or concomitant problems, degree of dependency of the patient, number of professional nursing care hours required per day, need for care by specialized nursing and support personnel, need for specialized equipment, frequency and methods of necessary surveillance of the patient, frequency and types of treatment procedures, and prognostic guidelines.

Operationalizing these concepts in terms of the specific services constituting hospital level care or patient conditions requiring hospital level care is more difficult given the differences in medical philosophy and practice that exist. Subjective judgments as to what constitutes appropriate acute level care have, in fact, been found to be fairly unreliable. In several early studies, reviewers using relatively unstructured subjective criteria achieved fairly high levels of total agreement. However, with the exception of Moorehead's (1964) study, such measurement methods were determined to be unreliable when corrected for the high rates of agreement occurring by chance. "Specific" agreement rates, which are the relevant standard, equal the ratio of the number of cases found inappropriate by both reviewers to the number of cases found inappropriate by either reviewer individually and so provide a more accurate assessment of reliability. Moorehead's (1964) study achieved an 88 percent overall agreement and 53 percent specific agreement rate. Other studies using subjective judgment have reported 73 to 88 percent overall agreement but much lower (0 to 40 percent) specific agreement. (Browning, 1965; Zimmer, 1967; Zimmer and Groomes, 1969; Zimmer, 1974.) It was found, however, by Zimmer (cited in 1967) that although agreement on which patients were over-utilizing was low, agreement on the percentage of over-utilization across all patients was high.

Various attempts have been made to improve the reliability of clinician judgment. Zimmer (1967) evaluated the possibility of improving observer agreement when data additional to that in the hospital record is made available, but found that increased information resulted in only slightly greater agreement between the paired observers. McClain (1972) incorporated a physician confidence scale in utilization review judgments and assessed its effect on the intra-reviewer and inter-reviewer reliability of physicians. He found that use of the scale predicted performance, but did not consistently increase intra- or inter-rater reliability. Rosser (1976) used a group consensus approach to evaluate the appropriateness of hospital utilization. She found interobserver agreement after the second round of Delphi analysis still manifested very wide variation, throwing doubt on studies based on judgments of one or two external observers, or of individual consultants. Similar results have been found with other dimensions of care. Richardson (1972), in evaluating the reliability of judgments on quality of care, concluded that to be useful retrospective evaluations should be based on pre-established standards for that care.

Based on such statistics, and the belief that subjective criteria were in part responsible for low levels of observer reliability (see Restuccia, 1976), attempts to develop explicit criteria were initiated. Early in the period, Donabedian (1969) outlined the basic requirements of an objective quality and utilization appraisal methodology. Because of the previously discussed historical differences in approach, however, criterion development was further advanced in the area of quality of care than that of utilization review. With respect to quality of care, a number of methodological approaches (see Williamson, 1979) had been proposed and refined, including diagnosis-specific criteria lists (Payne, 1965); critical indicators (Slee, 1974); trajectory methods which follow the patient over time assessing both the process of care and outcome (Brook and Stevenson, 1970), the use of tracer conditions (Kessner, Kalk & Singer, 1973) and the identification of critical events such as preventable adverse events or preventable progression of disease (American College of Surgeons and the American Surgical Association, 1976). The narrow scope of many of these approaches made them more suitable for research purposes than for general patient care monitoring; however, more comprehensive diagnosis-specific criteria sets (e.g., Payne, 1965; AMA, 1976) were available. In addition, many criteria sets had been integrated into methodologies specifically aimed toward quality assurance review by hospitals or PSROs such as the medical audit system developed by the Commission on Professional and Hospital Activities, (Ann Arbor, Michigan) or the Comprehensive Quality Assurance System (Rubin, 1975).

There were also some criterion-based methodologies available for utilization review at the time the project was initiated. Gonnella, Miller, and Smithline (1980), for example, outline an objective set of laboratory, radiological, pathology and treatment data collected on admission, the post-operative day, and day of discharge which, when aggregated, can highlight a range of patient care problems related to both quality of care and over-utilization of services. The manner in which these criteria might be used to create comprehensive guidelines for utilization review, however, was not well specified. Additionally, the American Medical Association (1976) in conjunction with 38 national medical societies developed a diagnosis-specific criteria set for use by hospital and PSRO review departments in admission certification, continued stay review, and medical care evaluation studies. The diagnoses in the set acccounted for 75% of the short stay admissions determined to be of concern in utilization review.

Two methods had been developed specifically for utilization review. (A third method, the Care Level and Timeliness Review (CLTR) developed by McKinsey and Company (Edwards, 1981) is a systematic procedure for establishing a utilization review system and does not incorporate any specific criteria for determining necessity.) The Appropriateness Evaluation Protocol or AEP (Gertman and Restuccia, 1981) was developed under contract to HCFA (SSA-600-75-0209) and was first introduced with 27 explicit criteria (covering both services and patient condition factors) for determining the appropriateness of a day. In the original version the appropriateness of the admission was determined subjectively. However, 18 admission criteria were added in a subsequent (1981) version, and minor changes were made to some of the days of care criteria.

The AEP is designed primarily as a retrospective review tool. In addition it does not follow individual patients over the course of the stay, but instead evaluates a cross-section of the patients hospitalized on a given day. The AEP also incorporates an "override" option to allow reviewers to indicate when a day not meeting any criterion should be judged appropriate or, conversely, when a day which does meet a criterion should be judged inappropriate. For each inappropriate day, reviewers are asked to select the most likely cause (e.g., delay in ordering, lack of availability of lower level care, etc.) from an itemized list.

To use the AEP, reviewers first memorize all criteria items and then examine the medical record, reading first the physician notes, then the nursing notes, lab reports, and social service notes for the day of review and the preceding day. The AEP data collection form is then completed and the actual evaluation made. The original AEP process required approximately 7 to 10 minutes for reviewing a day of care from an individual patient record. The more recent estimates of abstracting time following the revision of the AEP (to add admission criteria) were twenty cases per day or about 20 minutes on average for completion of a record.

The AEP achieved higher levels of reliability than those studies based on unstructured, subjective judgment discussed previously. For the first version specific agreement rates of 73% to 79% were found for the days of care review when evaluations of two reviewers were compared. The specific agreement rate for admissions was low, however, ranging from 31% to 40%. This was attributed to the subjective nature of the admission evaluation, a situation which was improved with the introduction of specific admission criteria. However, no assessments of the reliability of the recently-introduced criteria have as yet been published.

Another criterion-based instrument in use at the time the instrument was developed was Interqual's Intensity of Service, Severity of Illness, and Discharge (ISD) Screens (Jacobs and Lamprey, 1979). This system, designed for admission certification and concurrent review, incorporated over 200 criteria designed to assess medical necessity in terms of "the severity of the patient's condition, the intensity of treatment the patient is getting, and the absence of indications that discharge can be safely accomplished." The ISD criteria sets are organized into one generic and 13 body-system specific sets. If no generic criterion is determined to be applicable to a case, the most applicable body system criterion set is then examined. These criteria were developed by Interqual using a panel of physicians and were then used and reviewed by participating PSROs with subsequent modifications.

As of project initiation, this approach had not been systematically evaluated, although it had gained significant acceptance by PSROs and in-hospital utilization review programs. According to Interqual estimates, it was being used by 73 PSROs at the time of project initiation.

3. Design of the Preliminary Instrument

In the initial phase of the study few restrictions were placed upon the eventual form of the methodology. A number of alternatives were considered, including assessment of nonacute utilization (1) on-site at the hospital concurrently with the patient stay, (2) through retrospective examination of the medical record, and (3) through the use of automated data such as discharge abstract data.

At the outset of the study or early in its development, however, some of the essential features of the methodology were apparent. The intended use of the instrument, i.e., application in a large national survey, dictated a method that was not prohibitively expensive to apply. This, in turn, implied that the method should require a minimum amount of time per patient and that it not require extensive medical training for its use.

Reliability and validity of the method were, of course, key issues. Avoiding bias in stating the rate of nonacute utilization was of additional concern. Whereas some utilization review methods are designed to reject a substantial percentage of the cases which are subsequently reviewed by a physician advisor, it was felt that -- due to the potentially high cost of using such a procedure in a nationwide sample -- the instrument should minimize deliberate bias in either direction.

As noted, the project had at the outset rejected the use of length of stay norms to measure inappropriate use due to the analytical ambiguities inherent in this approach. Additional background research described above confirmed that the methodology should be based on explicit criteria, relying as little as possible on clinical judgment. Not only was this considered essential to ensure reliability, such a methodology was apt to require less highly trained medical personnel and so would be less costly to apply in a national survey.

The need for precise clinical criteria ruled out the use of automated data for the study, since no such data abstracting source had the data required by existing criteria sets for assessing inappropriate use, particularly for assessing the individual days of the stay. Of the two remaining alternatives -- concurrent on-site evaluation and retrospective evaluation using the medical record -- the former was rejected as being too costly given the magnitude of the survey to be performed. Following the example of the AEP, this study adopted retrospective record review as the basic methodology of the national survey, on the grounds that the function of the medical record is to document precisely that information on patient condition and services rendered throughout the stay which is also required to judge the necessity of the hospitalization and continued stay.

a. Components of the Preliminary Instrument.

Of the criterion-based methodologies identified in the background research phase of the study, two were considered applicable for use in the present study: the Appropriateness Evaluation Protocol or AEP (shown in Appendix B) and Interqual's Intensity of Service, Severity of Illness, and Discharge (ISD) Screens (shown in Appendix C). These together formed the basis of the present methodology. Although the AMA (1976) criteria sets were relatively comprehensive, the diagnosis-specific approach was considered too cumbersome to be applied in a widescale data collection effort. In addition, although the criteria were specific and detailed, they were often not defined precisely enough to be used in a consistent manner.

The AEP, as noted, consisted of criteria (18 related to service and 9 to patient condition) for judging the appropriateness of a day of stay for general medical and surgical patients. At the time of project initiation, the AEP lacked criteria for reviewing the need for admission. However, a revised version of the AEP, containing 18 admission criteria, was released midway through the instrument development phase of the project (see Gertman & Restuccia, 1981). As used by its developers, the AEP employed both RN and MD reviewers. In both versions, reviewers were allowed to override and, in addition, were asked to identify reasons (selected from a list of reasons itemized on the form) for each day judged to be inappropriate. The AEP was, in addition, a paper-and-pencil instrument with a structured application and decision process.

The ISD consisted of over 200 criteria organized into one generic and 13 body-system specific sets and was typically applied by a RN reviewer who referred questionable cases to a physician advisor. It also incorporated criterion sets, adopted from the American Association of PSRO's, for assessing the appropriateness of 13 major operative procedures. Three types of criteria were used in the ISD: those related to the severity of illness of the patient upon admission, those for the level of services provided on a given day, and those for the patient's readiness for discharge on a given day. For an admission to be considered appropriate, it had to meet a severity of illness criterion and a level of services criterion on the first day of the stay. A day was considered appropriate if it met a level of services criterion or if the patient did not meet a readiness for discharge criterion. Although the ISD contained an implicit logic for performing concurrent review, it did not incorporate a structured paper-and-pencil application process as did the AEP.

The instrument developed in the first phase of the study (shown in Appendix D) used the generic and body system-specific criteria sets of the ISD but developed a paper-and-pencil abstracting form similar to that of the AEP with specific instructions for abstracting information and evaluating the admission and days of the stay. The ISD severity of illness screens and surgery justification screens were used in evaluating the admission; the intensity of service and readiness for discharge screens were used in evaluating the individual days of care. The types of criteria and the manner in which they inter-relate to determine appropriateness are described in the following section. The instrument also incorporated a modified version of the "reasons" list used by the AEP.

The ISD criteria sets were selected over those of the AEP, first, because the AEP at the time had no admission criteria and, second, because they were more extensive than those of the AEP. Any criterion on the AEP that was not also on the ISD was added to the preliminary instrument. Given that the data collection needs of the national survey required a stand-alone instrument that could yield accurate estimates without the need for extensive clinical judgments, the completeness of the criteria sets were a primary concern. It was feared that without the fail-safe mechanism of physician review, the shorter criterion sets of the AEP, though easier to apply, might bias the estimated rates of inappropriate use upward by failing to include medically accepted reasons for hospitalization. The longer ISD criterion set could, on the other hand, be reduced or collapsed if the pretest phase showed that a number of criteria were too infrequently used to have an impact on measured rates of nonacute utilization.

b. Logic of the Preliminary Instrument.

A number of factors have been implicated as uniquely characteristic of acute hospital-level care, including the severity and stage of a patient's illness, his prognosis, the presence of complicating factors, the degree of functional impairment, and the services required by the patient. Acute-level services are specifically those which must be provided in the hospital environment, such as services requiring technologically advanced equipment or highly-trained personnel. In defining acute-level services it is recognized that some conditions do not necessarily require highly sophisticated diagnostic or therapeutic procedures but do need to be monitored in a closely-controlled environment. Thus, in many cases monitoring and observation are considered appropriate hospital-level services, as are the day or days required to prepare the patient for surgery.

Different criteria sets are relevant depending upon whether one is evaluating the necessity of admission to hospital level care or the necessity of a continued stay. For both, however, criteria can be classified essentially as either factors relating to patient condition or factors relating to level of service.

With respect to the necessity of the admission, two situations warranting the patient's admission to a hospital were distinguished: (1) those in which hospitalization is necessary because of some type of acute condition (usually requiring immediate medical intervention) or (2) those in which hospitalization is necessary because the patient has a problem (often longstanding) that requires services (e.g., performance of a diagnostic or therapeutic procedure) most reasonably performed on an inpatient basis but which would not require hospitalization in the absence of definitive treatment for that condition. "Elective" surgery is a special case of the latter category, which also includes such care as admission to an organized alcoholism rehabilitation program or medical treatment such as traction or organized physical therapy.

Ideally, for hospitalization to be appropriate both patient condition factors and level of service factors should be met, i.e.; the patient should be receiving acute level services and should actually have a condition (whether longstanding or in an acute or exacerbated phase) which requires those services.

Methodologies for determining the appropriateness of hospital level care differ in their reliance on patient condition and level of service factors as indicators of appropriate use. Many investigators (e.g., Goldberg and Holloway,

1975) have emphasized level of care exclusively as determining appropriateness, thereby assuming the need for such care. The admission criteria for the AEP (second version) incorporate both level of service criteria and patient condition criteria; either one alone is considered sufficient to justify admission to an acute care setting. The ISD as applied in concurrent review used a more stringent test: the patient was required to meet both a severity of illness criterion upon admission and a level of service criterion to warrant admission.

In designing the preliminary instrument for this study, patient need rather than services was considered the overriding factor indicating need for admission. Given patient need for hospital level care, the failure to provide the appropriate level of care was considered to reflect more on the timeliness of services (or quality of care), rather than the appropriateness of admission. Basing the justification for admission solely on services seemed to run the risk of justifying over-utilization on the basis of the utilization itself. That is, the logic is circular and "over-utilization" of a service could lead to a determination of an appropriate day. Therefore, the following strategy was adopted: The admission evaluation criterion set for the preliminary instrument would be comprised of the ISD severity of illness criteria, thereby basing the need for admission entirely on patient condition factors.

In the instrument revision process, however, we found it necessary to reintroduce level of services as a criterion for the justification of admission for some conditions, namely those for which admission is appropriate only for definitive treatment (e.g, intermittent claudication, herniated intervertebral disc). In this way the patient was required to meet both a condition criterion and an implicit level of service criterion in order for the admission to be judged as appropriate. The patient condition factor remained, however, as insurance against the misuse of services themselves to justify acute-care setting. The one exception to this principle was in the evaluation of admissions for elective surgery.

As noted, surgery justification sets were available for a handful of procedures, and these were included in the preliminary instrument. Reports from previous users of these sets indicated, however, that the criteria were so broad as to justify almost every procedure. Given that an extensive criterion development for surgery was beyond the scope of the study, it was recognized at the outset that we might not be able to evaluate the necessity of major elective procedures. As discussed below, these were dropped from the instrument following the preliminary evaluation phase. Since such procedures comprise a large proportion of hospital admissions, however, they were the one service allowed to stand as a criterion warranting admission in and of itself.

Similar considerations were applied in the development of criteria for assessing continued stay. Again the relevance of both patient condition factors and level of service factors must be considered. As in its admission criterion set, the AEP criteria for daily evaluation incorporated both types of criteria, allowing either to justify the appropriateness of a day. The ISD, on the other hand, required that the day both meet an intensity of service criterion and _not_ meet a readiness for discharge criterion to be considered an acute-care day. The methodology used in this study, both in its preliminary and revised version, retained this logic, requiring that both conditions be met on a single day for the day to be judged appropriate. (In the revised instrument, however, certain relatively intensive services (e.g., ICU/CCU care) or services that were part of an ongoing program (e.g., medication regulation, organized rehabilitation) were themselves allowed to stand as criteria indicating that the patient required continued hospitalization.)

In designing the preliminary instrument, the logic for evaluating days was taken to apply to every day in the stay, including the day of admission. Therefore, the _day_ of admission may be judged inappropriate even if the admission itself is judged appropriate. For example, a patient may be appropriately admitted for a cholescystectomy. If, however, through inefficient scheduling the patient waits three days for his surgery to be performed (assuming the patient has no other medical conditions which need care before surgery is possible), these two extra days -- including the day of admission -- can be considered unwarranted.

4. Evaluation of the Preliminary Instrument

One of the first concerns in the pretest instrument assessment phase was the validity of the criteria, with respect both to completeness and accuracy. Of particular concern were omissions in the criteria set. These are much more difficult to identify than criteria which are merely incorrect, and yet may seriously bias the results by over-representing inappropriate hospitalization (i.e., by labelling as inappropriate all cases which fail to meet an incomplete set of criteria). An intensive criterion-validation process was considered beyond the scope of the project. Since the ISD was in widespread use, however, the project had the opportunity to assess and modify the criteria sets based upon the changes that had been necessitated by practical use. To collect this information, PSROs using the ISD were surveyed to find out whether they had revised criteria sets available. Fifty-two organizations provided the project with revised criteria. In these sets, some criteria values had been modified and a number of criteria added; however, few deletions were found.

These modified criteria were compiled, with multiple versions of the same criterion placed together. The new sets were then reviewed by the study's Medical Consultant Panel (see Appendix A for a list of members). In a group process, the panel evaluated the criteria, resolving contradictions and making final determinations as to the inclusion or exclusion of criteria. In addition, the panel reviewed 52 medical charts that were to be used in the evaluation of the pretest instrument, rating the appropriateness of the admission and individual days of the stay.

The pretest instrument was then field tested in 15 hospitals, seven on the East Coast and eight on the West Coast. These 15 hospitals were evenly divided among major teaching, community teaching, and nonteaching hospitals. A primary concern in this effort was to assess the reliability and feasibility of the methodology that had been developed. Another objective was to identify the most reliable and cost effective method of using the instrument in terms of (1) the type of personnel for applying the instrument and (2) the method of training.

Specifically, we examined the performance of MDs, RNs, and medical records personnel in various review functions. In addition, mail and face-to-face training were compared. Previous studies (e.g., Dikewood Industries, Inc., 1979; Gertman and Restuccia, 1981) had found no differences between physicians and non-physician personnel in the reliability of applying explicit criteria. These studies, in addition, found that non-physician personnel including RNs (Gertman and Restuccia, 1981) and medical records technicians (Dikewood, 1979) formed judgments equivalent to those of physicians. The reliability of mail, as opposed to face-to-face training had not been assessed at the time of the study.

Instrument application functions were defined as (1) abstracting and (2) decision-making. Three combinations of personnel or "teams" were used to perform these functions: (1) ARTs as abstractors and MDs as decision-makers, (2) ARTs as abstractors and RNs as decision-makers, and (3) RNs who performed both abstracting and decision-making. Also, in a small subset of the sample, ARTs performed both abstracting and decision-making. Two sets of identically composed teams were assigned to each of the 15 hospitals. At the same time these teams reviewed the records previously rated by the Medical Consultant Panel, which were also independently evaluated by a sample of seven utilization review physicians, using their own judgment without the aid of the instrument.

Summary results are presented in Exhibits III-1, III-2, and III-3, which show the reliability of abstracting selected admission and daily criteria, and the reliability of evaluations of appropriateness. (Criteria were included in these tables based on their frequency of occurrence.) Exhibit III-1 shows high levels of agreement for abstracting admission criteria. Agreement overall (measured using tau which provides a close approximation to the kappa statistic) ranged from .77 to .89. Face-training produced consistently higher reliability coefficients than mail-training, but levels for mail-training were adequate, ranging from .69 to .84. ARTs, with scores ranging from .83 to .91, were more reliable abstractors than RNs (with scores from .69 to .84). Major teaching institutions produced by far the most reliable scores (probably due to the better documentation of their records) and nonteaching institutions produced the least reliable scores. It also appeared that the reliability of abstracting was

EXHIBIT III-1

PRETEST ANALYSIS

INTER-OBSERVER RELIABILITY OF MOST FREQUENTLY OCCURRING ADMISSION CRITERIA* (N=555)

CRITERION	OVERALL AGREEMENT	TRAINING TYPE:		ABSTRACTING TEAM:		HOSPITAL TYPE:			PATIENT AGE:		2nd DIAGNOSIS PRESENT:	
		FACE	MAIL	ART-ART	RN-RN	COMMUNITY	MAJOR	NON-TEACH	< 65	65 & OVER	YES	NO
TOXIC AGENTS IN BLOOD	.89	.96	.75	.90	.84	.68	1.0	.66	.89	.81	.88	.89
TOXIC AGENTS IN URINE	.84	.87	.80	.89	.77	.76	.92	.64	.82	.85	.84	.84
POSITIVE BACTERIA CULTURE	.81	.84	.72	.86	.72	.68	.92	.49	.83	.42	.78	.84
SIGHT LOSS	.77	.82	.69	.83	.69	.63	.87	.58	.78	.57	.73	.82
SPEECH LOSS	.83	.87	.76	.88	.75	.75	.89	.72	.82	.88	.84	.82
UNABLE TO MOVE	.82	.84	.80	.88	.74	.77	.89	.63	.84	.68	.83	.82
UNCONSCIOUSNESS/COMA	.81	.84	.75	.88	.70	.74	.89	.58	.83	.57	.76	.87
DISORIENTATION	.84	.88	.77	.91	.74	.73	.94	.63	.86	.61	.84	.85
GROSS HEMORRHAGING	.84	.88	.75	.89	.75	.70	.96	.58	.85	.61	.79	.91
WOUND DISRUPTION	.87	.88	.84	.91	.81	.81	.94	.67	.88	.76	.86	.88

* Kendall's Tau$_b$ is the measure of association. This statistic provides a close approximation to Kappa (the measure of first choice), which is not available in most statistical analysis program packages. Tau$_b$ was selected over two other common measures of association: Goodman & Kruskall's Gamma and Stuarts Tau$_c$. Gamma consistently overstates agreement for heavily skewed marginals, and Tau$_c$ is deceptively low relative to Kappa, especially with high incidence of agreement.

EXHIBIT III-2

PRETEST ANALYSIS

INTER-OBSERVER RELIABILITY OF MOST FREQUENTLY OCCURRING DAILY CRITERIA* (N=7354)

		TRAINING TYPE:		ABSTRACTING TEAM:		HOSPITAL TYPE:			PATIENT AGE:		MULTIPLE DIAGNOSES:	
CRITERIA:	OVERALL AGREEMENT	FACE	MAIL	ART-ART	RN-RN	COMMUNITY	MAJOR	NON-TEACH	LESS 65	OVER 65	NO 2nd DX	SECOND DX PRESENT
LEVEL OF CARE:												
MONITORING VITAL SIGNS	.54	.54	.52	.40	.77	.46	.69	.28	.53	.86	.51	.55
MONITORING BLOOD PRESSURE	.56	.57	.54	.42	.75	.49	.69	.42	.55	.52	.52	.58
MONITORING URINE OUTPUT	.64	.68	.52	.62	.70	.45	.78	.35	.65	--	.66	.63
MONITORING PUPIL REACTION TO LIGHT	.59	.58	.59	.47	.73	.52	.69	.43	.59	.54	.50	.63
MONITORING CENTRAL VENOUS PRESSURE	.56	.54	.59	.40	.72	.55	.69	.24	.55	.63	.49	.60
IV/IM MEDS	.71	.74	.65	.68	.80	.67	.80	.49	.71	.74	.67	.75
IV/MUSCULAR CHEMOTHERAPY	.72	.74	.65	.67	.81	.61	.83	.52	.71	.76	.67	.74
IV FLUIDS	.71	.74	.65	.66	.82	.60	.82	.52	.71	.74	.66	.74
OTHER MEDS REQUIRING 4HR MONITORING	.74	.76	.69	.70	.80	.63	.82	.52	.74	.59	.73	.74
VENTILATORY ASSISTANCE	.75	.76	.72	.75	.75	.66	.84	.47	.76	.64	.74	.76
CONTINUOUS 02	.75	.77	.72	.76	.75	.66	.83	.52	.76	.68	.76	.75
RADIOACTIVE MATERIAL IMPLANTATION	.71	.72	.68	.70	.74	.64	.79	.43	.72	.58	.68	.72
WHOLE BLOOD/BLOOD COMPONENT REPLACEMENT	.73	.75	.68	.72	.78	.65	.83	.46	.73	.73	.71	.74
EMERGENCY SURGERY	.70	.72	.63	.71	.63	.58	.79	.39	.71	.46	.74	.64
SURGERY REQUIRING GEN ANESTHESIA W/IN 24HRS	.69	.75	.54	.69	.71	.52	.82	.46	.69	.58	.70	.67
OBS OF CRANIOTOMY	.75	.79	.61	.78	.68	.50	.87	.39	.75	.67	.76	.72
OBS OF THORACOTOMY	.76	.79	.64	.78	.70	.57	.84	.42	.76	.637	.77	.73
OBS OF LAPAROTOMY	.70	.74	.56	.72	.66	.50	.80	.33	.71	.45	.73	.66
OBS OF COMPOUND FRACTURE REPAIR	.57	.62	.44	.59	.59	.35	.72	.34	.58	.47	.53	.60
OBS OF ORTHOPEDIC HIP SURGERY	.49	.53	.41	.46	.57	.30	.61	.38	.49	.46	.35	.59
OBS OF TRANSURETHRAL RESECTION	.57	.60	.49	.57	.60	.39	.70	.40	.57	.59	.48	.64
OBS OF PERIPHERAL VASCULAR SURGERY	.58	.61	.50	.61	.55	.42	.72	.28	.59	.47	.49	.64
OBS OF MYOCARD INFARCTION/ICU TRANSFER	.59	.63	.47	.63	.53	.37	.74	.26	.59	.45	.50	.63
OBS OF UNSTABLE PATIENT CONDITION	.37	.35	.42	.32	.48	.31	.45	.21	.37	.35	.23	.46
DISCHARGE CRITERIA:												
SCHEDULED PROCEDURE PERFORMED	.56	.55	.58	.63	.80	.36	.72	.23	.56	.52	.51	.59
TEMP BELOW 99 FOR LAST 24HRS	.51	.54	.45	.48	.66	.46	.71	.10	.52	.34	.44	.56
PRESCRIBED DIET WITHOUT NAUSEA	.58	.66	.43	.61	.58	.35	.79	.24	.59	.37	.52	.62
PASSING FLATUS/FECAL MATERIAL	.61	.68	.34	.66	.51	.30	.84	.27	.62	.43	.52	.67
VOIDING OR DRAINING URINE FOR 24HRS	.56	.64	.36	.63	.53	.25	.85	.11	.59	.25	.58	.55
TYPE/DOSAGE DRUG SAME FOR 2 DAYS	.61	.64	.54	.59	.67	.44	.72	.20	.63	.12	.57	.63
MEDICATION CONTROLLED PAIN FOR 2 DAYS	.53	.56	.48	.47	.64	.40	.72	.24	.54	.39	.49	.56
STEROID DOSAGE REGULATED FOR 3 DAYS	.74	.74	.74	.72	.77	.71	.84	.11	.74	.83	.77	.73
CHEMOTHERAPY PT HAS WBC OVER 2500	.59	--	--	--	72	74	49	45	59	--	67	--
WOUND CLOSING WITH NO NECROSIS INFECTION	.60	.61	.60	.57	.67	.56	.77	.24	.62	.35	.55	.65
NO DRAINAGE FROM WOUND	.63	.60	.66	.56	.72	.63	.75	.36	.63	.60	.57	.68
TOLERATING TUBE FEEDINGS FOR 2 DAYS	.57	.44	.72	.54	.60	--	.72	.48	.59	--	.49	.63
PT ABLE TO MANAGE ADEQUATELY	.60	.64	.52	.59	.63	.40	.72	.50	.61	.42	.58	.59

* Kendall's TAU_b is used as the measure of association

EXHIBIT III-3

PRETEST ANALYSIS

INTER-OBSERVER RELIABILITY OF EVALUATIONS OF THE APPROPRIATENESS OF ADMISSIONS AND DAYS OF CARE*

		TRAINING TYPE:		DECISION-MAKING TEAM:			HOSPITAL TYPE:			PATIENT AGE:		MULTIPLE DIAGNOSES:	
	OVERALL AGREEMENT	FACE	MAIL	RN-RN	ART-ART	MD-MD	COMMUNITY	MAJOR	NON-TEACHING	LESS 65	OVER 64	NO 2nd DX	2nd DX PRESENT
ADMISSION REVIEW (N=505)	.80	.78	.85	.76	.87	.96	.80	.85	.67	.82	.60	.87	.74
DAY OF CARE REVIEW (N=7004)	.81	.77	.88	.80	.83	.80	.84	.82	.73	.81	.70	.80	.82

*Kendall's TAU_b is used as the measure of association.

greater for patients under 65 years of age than over 65 and was more reliable for patients without a secondary diagnosis than those with. Because patients over 65 generally have multiple diagnoses, these two trends may be related.

Similar results are evident in Exhibit III-2 with respect to the reliability of abstracting daily criteria. The reliability coefficients for days, however, are generally lower than those for the admission and the trends less clear cut. The most obvious difference evident in Exhibit III-2 is that the reliability of abstracting in non-teaching hospitals is again substantially lower. The decreased reliability of abstracting daily criteria was attributed to the cumbersome nature of the daily review procedure, which required abstractors to fill out a new page for each day in the stay and to check every criterion that was met on a given day rather than stopping once a criterion was met, an extremely arduous process.

The reliability of evaluations of appropriateness -- shown in Exhibit III-3 -- were generally high: the reliability coefficient was .81 overall for days and .80 overall for the admission. Mail-training was actually somewhat superior to face-training although the difference was not significant. On the admission decision, MDs were most reliable, followed by ARTs. All types of reviewers were approximately equivalent in terms of the reliability of the day evaluations. Nonteaching hospitals again produced the least reliable results. Decisions made for patients over 65 or with a secondary diagnosis were also less reliable. The reliability of override decisions was lower than those based on the criteria, falling overall to .62.

With respect to the set of records that had been rated by the Medical Consultant Panel (MCP), field staff evaluations of inappropriate admissions and days were substantially lower than those of the MCP. The latter rated 11% of the admissions and 23% of the days in the sample of records inappropriate. Field staff found less than 5% of the admissions and 15% of the days to be inappropriate. The UR physicians in the sample exhibited rates similar to those of the field staff.

5. Instrument Revision Process

Based on the information obtained in the field test -- as well as practical considerations of feasibility and cost -- the instrument was substantially revised. The decreased reliability of abstracting days of care illustrated that the body system-specific approach was too cumbersome. As a result all criteria were combined into a single general set. This was facilitated by the fact that many of the same criteria or types of criteria were repeated throughout the body systems in the ISD criterion sets on which the instrument was based. In addition, very specific criteria were combined into more general categories as in "coma" (diabetic coma, alcoholic coma, hepatic coma) or "wound, open" which combined separate criteria from six body systems. There was no attempt, however, to devise a shortened list of criteria: all criteria on the original ISD sets were incorporated into the final instrument either explicitly or implicitly. This was because analysis did not indicate any of the criteria to be clearly redundant, although some criteria occurred much more frequently than others. In addition, in over a third of the cases encountered in the pretest phase, only one criterion was met, and this was just as likely to be a body system-specific criteria as a generic one. Deletion of more specific criteria, then, would seriously over-represent the proportion of inappropriate admissions or days.

In addition to the integration of the criteria, many vague criteria were re-worded and made more specific based on reviewer's comments, as well as on intensive scrutiny and discussion. Criteria that were not considered stringent enough were qualified with clearly phrased modifiers. This work was done with the advice and review of the study's medical and medical records consultants.

Other changes to the instrument included reversing the wording of the discharge screens so that only one criterion needed to be met to indicate the need for continued hospitalization (unlike the original ISD in which all discharge screens had to be met in order to indicate the inappropriateness of continued hospitalization). These were then renamed the "Need for Continued

Hospitalization" criteria. The reason list was collapsed on the basis of data showing that many of the more specific reasons could not be distinguished. (In addition, for the purposes of the present study, identifying very specific problems was not considered of high priority, the concern being more to identify broad classes of problems.) Finally, the logic and process of applying the instrument was considerably streamlined and the need for complex decision-making algorithms eliminated. Override judgments by reviewers were confined specifically to the need for admission, need for continued hospitalization, and level of services provided, rather than incorporating global judgments of appropriateness.

As a result of these changes, the need to separate the abstracting and decision-making functions was eliminated, so that a single individual could perform the whole procedure (making it a much less costly process). ARTs, rather than RNs, were selected for this function based on their superior abstracting reliability. The increased precision of the criteria and extensive training in the meaning of the criteria was considered methodologically superior to employing reviewers capable of sophisticated medical judgment.

Another major change at this point was the decision to eliminate surgical criteria. The 13 sets used in the pretest were too broad and tended to justify every surgery. Since a massive criterion-development effort was outside the scope of the present project, procedure or service-related admissions are evaluated only in terms of whether the procedure or service itself needs to be performed in an acute care setting. Operationally, this was done simply by referring to a list of procedures defined as appropriately performed in the hospital.

Lists of procedures defined as appropriate to be performed on an inpatient basis were taken from the Uniform Hospital Discharge Data Set (UHDDS) list of Class 1 procedures (U.S. National Committee on Vital Health Statistics, 1978). UHDDS defines Class 1, 2, and 3 procedures to be "significant" procedures which carry an operative or anesthetic risk or require highly trained personnel or special facilities or equipment. In general, Class 1 procedures correspond to surgeries, whereas Class 2 and 3 procedures are other significant diagnostic and therapeutic procedures. (Class 1 and 2 procedure lists can be found in Appendix E. Class 1 procedures are all ICD-9-CM procedure codes not covered by these lists.)

There are some Class 1 procedures which are routinely performed on an inpatient basis, but given special facilities such as ambulatory and same-day surgical units, may also be performed on an outpatient basis (e.g., D&Cs, bronchoscopies). Since medical practice is currently in a rapidly evolving state in this area, we did not choose to label these procedures as either appropriate or inappropriate for a hospital setting. These were instead given a special designation in the analysis so that their relative magnitude and impact could be estimated. In addition, to conserve resources, we made no attempt to compile a complete and exhaustive list of such procedures prior to the study. Rather, after the data collection effort was completed the actual procedures in the sample were examined to identify those potentially able to be performed on an outpatient basis. This process was carried out with the aid of the study's physician consultant. These procedures are so labelled in the analysis, and as medical practice changes, new estimates can be made by examining the detailed procedure codes contained in the data set.

Although many Class 2 procedures and some Class 3 procedures are routinely performed in an inpatient setting, the mere performance of these procedures was not considered sufficient to justify an admission. In almost every hospital stay at least one Class 2 or 3 procedure is performed. Allowing an admission on this basis was therefore considered too unrestrictive. It was felt that to be appropriate, an admission for a Class 2 or 3 procedure should be covered by one of the condition-related criteria given in the admission criterion set.

To increase the efficiency of the SMI, all conditions in the original ISD criterion sets used in developing the instrument were deleted if these conditions warranted hospitalization only for the performance of a procedure (e.g., senile cataract) and would not otherwise have required hospitalization.

6. Description of the Final Methodology

The instrument (named the Standardized Medreview Instrument or SMI) developed for the study employs a retrospective review of the medical record to assess whether hospital-level resources were used in the optimum manner. As used in this study, the instrument is applied by trained medical records professionals (Accredited Records Technicians, or ARTs, and Registered Records Administrators, or RRAs) using a set of criteria and a standardized application procedure to identify instances of unnecessary or inefficient hospital utilization. The instrument divides the evaluation of utilization into two phases: (1) evaluation of the admission and (2) evaluation of continued stay. The latter, in turn, examines two aspects of the stay: (1) whether continued hospitalization is warranted by the patient's condition and (2) whether acute level services are actually being provided the patient. Both conditions must be met for the day to be judged appropriate. The instrument incorporates 117 criteria for judging the appropriateness of the admission, 30 for evaluating the level of service and 26 for evaluating the need for continued hospitalization.

The Standardized Medreview Instrument (found in Appendix F) consists of a patient background section (Page 1), the utilization review section (Pages 2-9), reason and phase of stay codes (Page 10), comments section (Page 11), and sampling instructions (Page 12). The patient background section was used to abstract relevant information on the patient stay (e.g., age, sex, admission and discharge date, discharge status, diagnoses, procedures). The utilization review section -- which comprises the actual methodology for determining the appropriateness of utilization -- has two components: the admission review section (Pages 2-5) and the review of the individual days of care, including the level of service evaluation (Pages 6-7) and the need for continued hospitalization (Pages 8-9).

To apply the instrument, the reviewer first completes the patient background form, and then reviews the record for information necessary to perform the admission and daily evaluations. The Admission Review section is then completed in the following manner. The reviewer first determines whether any of the 117 criteria for admission are met. The criteria for admission are divided into four sub-sections: (A) Laboratory Abnormalities, (B) Medical Problems/Diseases/Complications of Surgical or Medical Care, (C) Abnormal Signs, and (D) Symptoms. These sections are ordered from the most objective findings to the least objective, and the reviewer is instructed to consider them in that order. For example, if a patient has both abnormal laboratory findings (Section A) and abnormal signs (Section C), the reviewer records laboratory abnormalities as the criterion for admission.

Following this, the reviewer completes the Admission Evaluation Section (shown in Exhibit III-4). An admission can be evaluated as justified on the following grounds: (1) one of the admission criteria is met, (2) no criterion is met, but the admission is for a Class 1 operation or procedure which requires the acute care setting, (3) no criterion is met, but the need for admission is supported by the logical extension of one of the criteria or (4) no criterion or logical extension of a criterion applies, but the admission appears justified on other grounds. The last two categories essentially provide options for the reviewer to override the criterion set.

An admission can similarly be judged as unjustified if (1) no criterion is met, or (2) a criterion is met but the need for admission is not supported on other grounds.

The next two sections of the instrument for Review of Individual Days of Care allowed for assessment of up to seven days of a stay. (For hospital stays longer than 7 days, the days to be abstracted were determined by a random process.) The determination was based on the Level of Services being provided to the patient on each day and the Need for Continued Hospitalization on that day. The SMI requires the abstractor to evaluate the level of service and the need for continued hospitalization independently to increase the reliability of

EXHIBIT III-4

DECISION PROCESS FOR DETERMINING THE APPROPRIATENESS OF ADMISSION

II. ADMISSION EVALUATION (check one):

APPEARS JUSTIFIED BY CRITERIA AND/OR RECORD REVIEW
(11-14)

☐ = 1000 Criterion met

☐ = 2000 No criterion met, but admission was for operation/procedure which requires acute hospital setting (see Manual Appendix A)

☐ = 3____ No criterion met, but need for admission supported by logical extension of criterion #_____. (Explain below)

☐ = 4000 No criterion met, but admission appears justified on the basis of record review. (Explain below)

APPEARS NOT JUSTIFIED. (Requires Reason Code from list 1) →

REASON CODE		
MOST LIKELY	OTHERS (OPTIONAL)	
(15-16)	(17-18)	(19-20)

☐ = 5000 No criterion met.

☐ = 6000 Criterion met, but record review does not support need. (Explain below)

EXPLANATION ____________________________________

____

____

DO NOT WRITE BELOW THIS LINE

(21-80)

CARD 3
BEG

1/3

2-5/________

6-7/____

8-10/______

11/14________

15-16/____

17-18/____

19-20/____

21-80/see box

CARD 4
END

TURN PAGE FOR REVIEW OF INDIVIDUAL DAYS OF CARE: LEVEL OF SERVICE

the decision. The final determination of the appropriateness of the day made subsequently in the analysis is based on the combination of the level of service and need for continued hospitalization reviews, thereby allowing the different types of inappropriate utilization to be distinguished in the analysis.

These sections, like the Admission Review section, first present the criteria themselves and then provide a section for making the evaluation. As with the Admission Evaluation, the level of services and need for continued hospitalization can be judged appropriate because (1) a criterion is met, (2) an extension of a criterion is met, or (3) no criterion applies, but the record review indicates that acute care was delivered/required. A negative decision can similarly be made on the basis that (1) no criterion is met or (2) a criterion is met, but record review does not support its use.

For any admission or day of care evaluated as nonacute the reviewer is required to indicate the most likely reason (and, optionally, up to two additional reasons). If the admission or need for continued hospitalization is judged to be unjustified the reason is chosen from List 1; if the level of service is nonacute, the reason is chosen from List 2 (both lists are shown in Exhibit III-5). List 1, for evaluating the need for admission or continued hospitalization, requires the reviewer to distinguish between patients who require institutional care at a subacute level and those for whom no further institutional care of any kind is required.

The reviewer is also required to indicate (at the bottom of the page for evaluating level of service) the phase of stay for each day evaluated. The phases are defined as pretreatment phases (workup in progress; treatment scheduled), treatment phases (surgical; medical; transitional) and post treatment phases (treatment completed). Definitions for each of these phases (shown in Exhibit III-6) are included in the SMI.

EXHIBIT III-5

REASON LIST USED TO SPECIFY CAUSES FOR UNNECESSARY ADMISSIONS AND DAYS OF CARE

List 1: ADMISSION OR CONTINUED HOSPITALIZATION APPEARS NOT JUSTIFIED

PATIENT REQUIRES NO INSTITUTIONAL CARE, BUT IS ADMITTED OR RETAINED BECAUSE:

11. Terminal patient for humanitarian reasons
12. Patient, family, physician prefer/insist on acute care. Includes non-compliant patient (*refuses* to comply)
13. Delay in orders or arrangements for discharge
14. Delay in discharge pending completion of patient or family education for home care
15. No home or alternate care is possible or available. (Includes patients who *cannot* comply, hostile family situation, or adverse home or environmental factors)
16. Other. Explain in Notes.
17. No reason can be deduced from available information

PATIENT REQUIRES SUBACUTE LEVEL OF CARE, BUT IS ADMITTED OR RETAINED BECAUSE:

21. Terminal patient for humanitarian reasons
22. Condition improving, recovery anticipated within 3 days
23. No lower level of care available: facilities will not accept patient or no beds available
24. Difficulty in placing patient in lower level facility. Includes Financial and Medical Difficulties, as:
    - Awaiting medical clearance from alternative facility
    - Awaiting financial clearance from alternative facility
    - Insurance (or Medicare/Medicaid) coverage is broader in an acute care setting, or 3 day Medicare qualification is needed
25. Difficulty in arranging transfer to other facility or delayed initiation of discharge planning
26. Patient, family or physician prefers or insists on acute level hospitalization. Includes non-compliant patient, teaching or research reasons, and legal reasons (controversial patient; psychiatric requirements; real, implied or feared threat of malpractice).
27. Other. Explain in Notes
28. No Reason can be deduced from available information

EXHIBIT III- 5 (Cont'd)

REASON LIST USED TO SPECIFY CAUSES FOR
UNNECESSARY ADMISSIONS AND DAYS OF CARE

LIST 2: LEVEL OF SERVICE APPEARS TO BE NONACUTE

DELAYS IN SCHEDULED SERVICE DUE TO:

31. Tests or consultations not complete or results not yet available
32. Change in patient condition
33. Patient or family noncompliant or indecisive regarding treatment course
34. Resource availability precludes timely delivery of service. Includes both physician and hospital causes, as, hours of operation (40 hour work week), staff shortages, emergency situations which may "bump" patient, or scheduling conflicts or delays, including transfers to other facilities.
35. Sequencing conflicts or delayed ordering by hospital or physician. Includes premature admission (e.g., Friday admission to hold bed for Monday surgery, sequencing of tests causes extended stay, etc.)
40. Patient does not require acute level services (i.e., patient did not need to be admitted or patient is ready for discharge).
50. Other. Explain in Notes.
60. No reason can be deduced from available information.

EXHIBIT III-6

PHASES OF STAY: DEFINITIONS

PHASE OF STAY

PRE-TREATMENT PHASES:

1. Diagnostic work-up in progress (prior to determination of definitive therapeutic course).
   NOTE: Phase 1 applies only when changes in the patient condition or diagnosis indicates the patient is in an "evolving" status. Once a therapeutic course has been determined, use Phase 2 if no therapy has begun, or Phase 3, 4, or 5 if therapy is underway.
2. Scheduled procedure, treatment, or management pending (prior to initiation of definitive therapeutic course).
   NOTE: Phase 2 may be used for patients entering hospital whose treatment course has been determined prior to admission.

TREATMENT PHASES:

3. Surgical — includes day of procedure and post-op observation period; as defined in Level of Service criteria.
4. Medical Therapy — includes drug, respiratory or physical therapy, psychotherapy, or traction.
5. Transition from one therapeutic regimen to another: Use when any therapy is being given, but focus is changing. If significant complications develop, or entire course of treatment is being re-evaluated use Phase 1.

POST-TREATMENT PHASE:

6. Planned course of therapy completed.

B. DESIGN OF THE NATIONAL SURVEY

1. Development of the Sampling Plan

The goals of the study called for the evaluation of both admissions and days of care. Accordingly the sampling plan for the pilot study used a three-stage cluster sample of days of care: the first stage sample, i.e., primary sampling units (PSUs), were acute-care, short-stay hospitals, the second consisted of records within each sampled hospital and the third consisted of days within a stay. The sample was designed as follows:

a. Determination of Optimum Sample Sizes

The first step in designing the sample was to determine the optimum sample size given study resources, and the optimum distribution of the sample over PSUs, admissions, and days of care. Precision is generally reduced in a cluster sample relative to that obtainable in a simple random sample with the same number of observations. This effect increases as clusters become more similar, i.e., as the days in a given stay become more similar to each other and less similar to other days that might have been sampled and as the admissions in a given hospital are more like each other and less like admissions in other hospitals. The clustering effect is measured by δ , the measure of homogeneity (or intra-cluster coefficient).

The following formula expresses the precision of a cluster sample in terms of that obtainable for a simple random sample with the same units of observation.

$$\sigma_C^2 = \sigma_{SRS}^2 \left[\delta_1(\bar{n}\bar{\bar{q}} + 1) + \delta_2(\bar{\bar{q}} - 1)\right]$$

Where:

σ_C^2 = the variance of the estimate in the cluster sample

σ_{SRS}^2 = the variance of the estimate in a simple random sample of the same size (e.g., of size $m\bar{n}\bar{\bar{q}}$).

m = the number of hospitals

$\bar{n}$ = the average number of records per hospital

$\bar{\bar{q}}$ = the average number of days per record

δ^1 = the measure of homogeneity of nonacute days within hospitals

δ^2 = the measure of homogeneity of nonacute days within records

As can be seen, as homogeneity is increased, greater precision is achieved by sampling fewer days from more records and fewer records from more hospitals. The trade-off in terms of study resources between increasing the number of hospitals versus increasing the number of records per hospital (and similarly the number of records and the number of days per record) depends, not only on cluster homogeneity, but the relative cost of collecting data at the three levels. The following formulas are applicable:

$$\bar{\bar{q}} = \sqrt{\frac{1 - \delta_2}{\delta_2} \; \frac{C_2}{C_3}}$$

$$\bar{n} = \frac{1}{\bar{\bar{q}}} \sqrt{\frac{1 - \delta_2}{\delta_1} \; \frac{C_1}{C_3}}$$

$$m = \frac{C}{C_1 + C_2\bar{n} + C_3\bar{n}\bar{\bar{q}}}$$

Where: $\bar{\bar{q}}$, $\bar{n}$, m, δ_1, and δ_2 are defined as above and

C = total survey funds available
C_1 = per hospital costs
C_2 = per record costs
C_3 = costs per day per record

Based on a hypothetical distribution of inappropriate days and admissions over hospitals and records δ_1 was estimated at .20 and δ_2 was estimated at .005. C_2, the cost of abstracting a record, was estimated to be 10 times higher than C_3 , the cost of abstracting a single day within a record. It was estimated that given the same amount of time the abstractor could abstract ten days in one record or one day in ten records each). Field study costs were identified in detail and broken down into fixed costs per hospital (e.g., travel to and from the hospital) and costs dependent upon the number of records abstracted (e.g., reimbursement for abstractors' time). The ratio of per-hospital costs (C_1) to per-record costs (C_2) was estimated to be about fifteen to one.

Using these formulas and values, the optimum sample sizes (given project resources) were determined to consist of approximately 240 hospitals with 20 records per hospital and six days per record. (Because certain hospital types were oversampled, the final sample drawn for the study was 255 hospitals, a figure which is in the breadth of the optimum. In addition, seven days, also within the breadth of the optimum, rather than six days per stay were abstracted.)

b. Development of the Sampling Frame

Only nonfederal, short-term, general hospitals were included in the sample. Since pediatric cases were excluded from the study, pediatric hospitals were also excluded. In addition, no hospitals were sampled in the U.S. territories or in Alaska and Hawaii.

The population of interest was identified by the 1980 National Master Facility Inventory of Hospitals. Exhibit III-7 presents the distribution of institutions covered in the survey by type of service. Only hospitals classified as "general medical and surgical" (Code 10 under "type of service") and not classified as federal (Codes 41 through 48 under "hospital classification") were included in the sampling frame. Hospitals with a region code of 0, for territories, or a state code of 94 or 95, for Alaska and Hawaii, were also excluded. As shown in Exhibit III-7, there were 5,696 nonfederal general medical and surgical hospitals in the continental U.S. in 1980. Of these, 8 were A.H.A. registered long-term care hospitals and so were excluded from the sampling frame, leaving a total of 5,688 hospitals covered by the sampling frame.

c. Development of the Strata and Sampling Strategy

One of the goals of the sample design was to ensure adequate representation of certain hospital characteristics of interest. This is generally accomplished by stratifying on variables for which representation is desired and possibly oversampling in some strata (although, oversampling should be minimized since -- while increasing precision within the oversampled subpopulation -- it decreases precision of the estimates overall). The following stratification variables were considered, since these characteristics were felt to be particularly relevant to the classification of hospitals and to the possible explanation of inappropriate utilization:

- Region (4 categories)
- Ownership (3 categories)
- Bedsize (5 categories)
- Teaching Status (2 categories)
- Urban/Rural Location (2 categories)
- Occupancy Rate (4 categories)

EXHIBIT III-7

DISTRIBUTION OF HOSPITALS IN THE 1980 AHA SURVEY BY TYPE OF SERVICE
(excluding territories and Alaska and Hawaii)

	Total	Short-Term
General Medical and Surgical		
Nonfederal	5,696	5,688
Federal	310	301
Hospital Unit of an Institution	73	73
Hospital Unit of an Institution for Mentally Retarded	15	15
Psychiatric	467	151
Tuberculosis and Other Respiratory Disease	10	1
Obstetrics and Gynecology	14	14
Eye, Ear, Nose and Throat	19	19
Rehabilitation	63	18
Orthopedic	9	9
Chronic Disease	52	1
Other Specialty	95	50
Institution for the Mentally Retarded	57	1
Alcoholism and Other Chemical Dependency	56	51
Children's General Medical and Surgical	51	50
Children's Hospital Unit of an Institution	1	1
Children's Psychiatric	35	34
Children's Tuberculosis and Other Respiratory Disease	0	0
Children's Eye, Ear, Nose and Throat	0	0
Children's Rehabilitation	9	8
Children's Orthopedic	22	18
Children's Chronic Disease	1	0
Children's Other Specialty	15	13
	7,070	6,516

In addition, because of their strong financial incentives to control unnecessary utilization, HMO-owned or affiliated hospitals (although representing only a small group of hospitals) were considered of special theoretical interest.

The basic stratification variables used in the design were region, ownership, and bedsize, with teaching status, urban/rural location, and occupancy rate used in a subsequent controlled selection process. Bedsize was selected not only for its substantive interest, but because (as will be discussed below), it was needed in the control of variation in the size of clusters. In addition, HMO-affiliation was included as a stratification variable since the small numbers of hospitals in this category necessitated oversampling in order to ensure a sufficient sample size for analysis. The stratification process was performed separately for the HMO hospitals and the remaining hospitals in the sampling frame.

i. Development of Strata for the Non-HMO Hospitals

The process of selecting the basic strata for this set was as follows. The thirty HMO-affiliated hospitals (to be discussed below) were first removed from the sampling frame. Three variables were then chosen for stratifying the remaining hospitals. These were region (Northeast, North Central, South, West); ownership (nonfederal government, nonprofit, proprietary); and bedsize (0-99; 100-249; 250-499; 500-1000; 1000+). These variables were completely crossed. Five categories of bedsize were used in order to obtain relatively fine groupings in terms of the measure of size (MOS), which in this case was chosen to be the number of annual admissions. These categories yielded 60 strata overall. Some of these strata had too few admissions to comprise their own sampling strata. Therefore, new strata were devised by collapsing across bedsize categories within region-ownership categories to create basic strata. In some cases, new cut-points for bedsize were used in order to devise more equivalent strata in terms of aggregate admissions. Proprietary hospitals were found to represent only 9% of total acute-care hospital admissions, which with probability-proportional-to-size sampling would produce only 22 profit hospitals in a sample of 240. The number of strata of proprietary hospitals was increased in order to improve precision for this group, with the goal of producing 195 nonprofit and government hospitals and 45 proprietary hospitals in all.

In cluster sampling, one danger is that extreme variability in the size of clusters may have a serious effect on the variance of the estimates. Given the wide variability that exists in both the number of admissions and days of care per hospital, the sampling strategy sought ways to minimize the contribution of cluster size variation to the overall sampling variance. Two widely used methods for minimizing this effect include (1) sampling the primary selection units, i.e., hospitals, with probability proportional to size and (2) stratifying PSUs into groups of equivalent size prior to selection. These two approaches were employed in the present survey. Within the 42 region by ownership by bedsize strata identified above, 240 individual strata with approximately equal numbers of admissions per stratum were devised with the goal of sampling one hospital per stratum with probability proportional to size.

Because of the increase in the number of proprietary strata, the process of devising strata was done separately for the proprietary versus the remaining strata. With HMO-affiliated hospitals excluded, proprietary hospitals accounted for 3,132,085 admissions in 1980 and nonprofit/government hospitals accounted for 32,023,787 admissions. The proprietary hospitals were divided into 45 strata, each representing approximately 69,000 admissions, whereas the nonprofit/government hospitals were divided into 195 strata each representing approximately 164,000 admissions. These strata were constructed within the region, ownership, and bedsize categories identified previously; that is, no strata were devised which combined two regions or two ownership categories, and all strata represented the same or contiguous bedsizes. Whenever possible, the strata were devised so as to be pure with respect to teaching status, urban/rural location, and occupancy rate (i.e., to represent only hospitals with the same teaching status, urban/rural location, or occupancy rate) in order to increase the representativeness of these subpopulations in the sample.

Once these 240 strata were defined, a controlled selection sampling procedure was applied using (1) a four-way cross of teaching status and urban-rural location and (2) a five-level breakdown of occupancy rate (i.e.,< 40%, 40% to 59%, 60% to 79%, 80% to 89%, 90% +) as the control variables. Controlled selection provides a way to increase the representativeness of small samples by increasing the likelihood that the distribution of sampled units will correspond to the marginal distribution of the control variables. The actual process of applying the control selection procedure was carried out at the Institute of Social Research/University of Michigan using computer software developed there (documented in Hess, Riedel, and Fitzpatrick, 1975). Once the control selection pattern was selected, hospitals were drawn from the strata with probability proportional to size. This probability was equal to the number of admissions for the stratum as a whole. The controlled selection procedure resulted in multiple selections from 10 strata at the expense of 10 others. One additional hospital was sampled from each of these 10 empty strata for a total sample size of 250 hospitals.

ii. Developing Strata for the HMO Hospitals

The process of sampling HMO-affiliated hospitals was carried out independently. The first step in this process was the definition of the sampling frame. This group of hospitals, while of theoretical interest, is a difficult one to define and identify. Of sole interest are hospitals which have strong affiliations with an HMO and are either managed by them or admit primarily HMO patients, since these are the only hospitals in which the HMO could be expected to exert any marked control on utilization. Very few hospitals are actually owned and managed by HMOs since only the largest plans have the resources to establish and maintain hospitals. There are more hospitals which are not owned by HMOs but have formal relationships with them and whose clientele consists primarily of HMO patients. However, these are not readily known.

The data tapes for the 1980 American Hospital Association Annual Survey of Hospitals includes a variable pertaining to HMO affiliation. It was found, however, that this category consists not only of hospitals with formal relationships with HMOs but a variety of informal relationships as well (having been elicited by the question: "Does your hospital provide treatment to members of an HMO on another basis than emergency" which was too broadly worded to exclude loose and informal associations). InterStudy of Minnesota was also contacted in the effort to identify the population of interest. This organization had no data on the percentage of admissions in any given hospital that were HMO enrollees or even a listing of hospitals with strong HMO affiliations. An informal survey had been done to determine which HMOs had been sponsored by hospitals: it was not necessarily the case, however, that a hospital sponsoring an HMO will draw all or even a significant percentage of its patients from the HMO.

The most reliable information on HMO-affiliation was obtained from a survey of HMOs conducted for the National Center for Health Services Research in 1980 (Contract No. 233-80-3004). This survey identified all HMOs that own their own hospitals -- and thereby excludes HMOs with a strong formal affiliation short of ownership. Thirty hospitals were identified in this manner, all but three of which belong to the same organization. Although this population excluded certain HMO-affiliated hospitals of interest it was felt at least to represent a relatively well-defined population. It was decided, therefore, to sample 15 hospitals from this group. As with the profit and nonprofit/government sample, strata which were of equal-size with respect to number of admissions were established. The 30 hospitals had a total of 271,369 admissions, which when divided into 15 strata, yielded an average stratum size of approximately 18,000 admissions. Eleven of the hospitals were sufficiently large (having overall admissions equal to or greater than the average stratum size) that they were included in the sample with certainty. The remaining 19 hospitals were divided into four equal-sized strata (with respect to aggregate admissions). One hospital was sampled from each of these strata with probability proportional to size.

d. Sampling of Admissions and Days

Records were sampled only for patients discharged in 1981. In the sampling of records all cases had an equal probability of selection, e.g., there was no disproportionate sampling by diagnosis or body system. However, both pediatric cases and normal deliveries were excluded from the population. Pediatric cases were defined as any patient less than 16 years old or any patient discharged from a pediatric service. Normal delivery was defined as any delivery without complications (ICD-9-CM diagnostic code 650). Sampling within hospitals was done using a strategy designed to produce a self-weighting sample. An overall sampling rate designed to produce the desired number of records across the sample as a whole was established. THis sampling rate established the overall probability of selecting a record which took into account the expected number of pediatric cases, normal deliveries, and missing or incomplete records. Separate sampling rates were established for the nonprofit/government hospitals, the for-profit hospitals, and the HMO-affiliated hospitals. The within-sampling rate (WSR) was then established as:

$$\text{WSR} = \frac{\text{stratum MOS}}{\text{hospital MOS}} \times \text{overall sampling rate}$$

where MOS is the measure of size (in this case aggregate admissions).

For example, the overall sampling rate for proprietary hospitals was set at one record per every 1731, a rate which would yield a total sample (before excluding missing and out of scope records) of 1809 records from these hospitals (i.e., 3,132,085 total admissions divided by 1731). The within sampling rate for a hospital of 11,352 admissions drawn from a stratum containing 79,349 total admissions was 1/248 or

$$\frac{79{,}349}{11{,}352} \times \frac{1}{1{,}731}$$

Records were selected from the discharge records of the Medical Record Department using a systematic sampling procedure, by which every nth discharge was selected (after a random start). For hospitals with a small number of admissions or with a numbered discharge record the sampling was sequential throughout the year. For hospitals with large numbers of admissions and no sequentially-numbered discharge record, the sampling procedure first selected a systematic sample of days (approximately one in ten) and then a systematic sample of records within a day. This was done in such a way that the sampling rate (i.e., the rate for sampling days times the rate for sampling records within a day) was equal to the original WSR. The medical record numbers corresponding to the selected discharges were identified by Medical Records Department (MRD) personnel, who then pulled the charts for the abstractor. MRD personnel were not required to pull the charts for normal deliveries, pediatric cases, or newborns (if these were readily apparent) but had to indicate when this occurred. An example of the straight sequential and the sequential-within-days sampling plans generated by the project and sent to hospitals is given in Appendix G.

Once the abstractor obtained the sampled charts, he or she randomly ordered them, using the "Abstracting Record" provided for each hospital (see Appendix H). The abstractor then began with charts at the top of the random list and continued abstracting until all records had been reviewed or until the maximum number of charts desired had been abstracted. The abstractor did not review charts which were incomplete or had been removed by the Medical Record Department and also excluded any newborns, normal deliveries, or pediatric cases found in the sample. Records for patients not discharged in 1981 or who occupied long-term care or intermediate-care beds were also excluded.

Once a record was selected for abstracting, the abstractor reviewed up to 7 days within the stay. The day of discharge was always excluded, unless the patient was admitted and discharged on the same day. The sampling rate for days was set at one (i.e., all days were abstracted) if the length of stay (LOS) was seven days or less. It was set at one-half for lengths of stay between 8 and 14 days, at one-third for lengths of stay between 15 and 21 days, and so forth.

The initial day to be abstracted was selected randomly, with all days (except the day of discharge) having an equal probability of selection. The remaining days abstracted were those immediately following. If the number of days to be abstracted was greater than the days remaining in the stay, the abstracting continued with the first day in the stay. The sampling plan for days was computer-generated by project staff and attached to the back of each instrument.

2. Data Collection

The data collection effort was carried out nationwide and coordinated from SysteMetrics' Santa Barbara office. Since field test results had established mail training to be as reliable as face-to-face training, the recruiting and training of abstractors was carried out at long distance through use of telephone and mail communication. Hospitals were also recruited at a distance, and no site visits were performed to obtain hospital participation.

a. Hospital Recruitment

After hospitals were sampled, the data collection phase was initiated with the hospital recruitment effort. In order to increase the response rate, the American Hospital Association (AHA) and state and local hospital associations were contacted and asked for their endorsement. After thorough consideration, the AHA approved the project and provided a letter of endorsement addressed to sampled hospitals. In addition, HCFA regional offices and PSROs were notified of the study by the Health Standards and Quality Bureau to encourage participation.

Hospitals were contacted first by mail, with a subsequent telephone follow-up. If a hospital refused to participate, a "shadow" hospital was selected from the same stratum (again with probability proportional to size). Additional shadows were selected as necessary until all strata were represented.

b. Recruitment of Abstractors

Abstractors were initially recruited through advertisements in the Journal of the American Medical Record Association and in the newsletter of AMRA, "Counterpoint," as well as by contacting abstractors known to project staff from previous data collection projects. Later, for geographic areas from which no response was obtained, various organizations were contacted, including PSROs, colleges and universities with Medical Record programs, state and local Medical Record Associations, and finally, the hospitals participating in the study. In a total of 19 hospitals, abstracting was performed by individuals currently or formerly employed by the hospital. There were no significant differences found between this group and the remaining hospitals in the rates of nonacute utilization found.

Over 150 potential abstractors were considered for this project. Initial screening of abstractors involved reviewing resumes, followed by a telephone interview. Experience in Utilization Review (UR) and/or a good understanding of the processes involved in UR were the prime qualifications desired for abstractors, and each abstractor employed on this project possessed either an A.R.T. (Accredited Record Technician) or an R.R.A. (Registered Record Administrator) degree. A secondary consideration in choosing abstractors was their availability, since many were working full time and could only abstract on weekends. These factors were coordinated with the requirements of participating hospitals as to when access to records would be allowed. Also considered was geographic proximity of the abstractor to the hospitals. To keep costs within reason, long distance travel was minimized. However, when no qualified abstractor was available in the vicinity, long-distance travel was elected.

c. Training of Abstractors

Prior to beginning work in any hospital each abstractor was required to spend one day working independently with the self-instructional training materials. This activity consisted of reading an extensive training manual, followed by the completion of five abstracts using records that had previously been rated by the Medical Consultant Panel. The five records were chosen with different characteristics to provide a broad range of criteria and evaluations. They included:

1. A cardiac patient with a three day ICU stay.
2. An elective surgery.
3. An emergency surgery.
4. An infectious condition with complications of sepsis, anemia, and dehydration.
5. An inappropriate admission for work-up of arthritis.

The training manual (SysteMetrics, 1983) had three sections:

- Background of the project
- Appendices (listing Class 1 and Class 2 procedures and major drugs by type).
- Logic, Field Staff Instructions and Details of the Instrument

The instructions for use of the instrument included detailed discussion of the individual criteria, as well as a description of the correct procedure for application of the instrument. All abstractors were told to call at any point in their training (as well as during the actual abstracting in hospitals) for clarification on use of the instrument, or for advice on particular cases that were being abstracted. A WATS line with a toll free number was available for their use.

Once completed, the five abstracts were returned to our staff and reviewed. The abstracts were coded and checked against a master set, based on the MCP evaluations. Discrepancies were noted, and each abstractor received another telephone call, during which he or she was informed of the types of errors made. Variations were discussed, as well as appropriate application of certain criteria, and the logic of the instrument was reinforced.

The most significant matter to be discussed with the abstractors following training was their understanding of the concept that calling an admission or a day not justified or non-acute did not necessarily mean that the patient did not need care, but rather that the care could have been provided at a facility other than an acute care hospital. Other discrepancies that occurred frequently included:

- Misinterpretation of pay source.
- Listing other than Class 1 procedures in the Background section of the instrument.
- Conversion to military time.
- Use of specific criteria, (e.g., inaccurate lab values taken from Discharge summary without verification by actual lab report, or identification of an organized Rehabilitation program versus occasional Physical Therapy services).

Of course, not all abstractors made all of these errors, and in fact most made only two or three types of errors throughout all five charts.

Following submission of the training exercises, six abstractors were found to be unsatisfactory and were not employed in the study. Another abstractor was required to complete an additional set of training exercises after which his performance was completely satisfactory. A total of 78 abstractors were used for the 253 hospitals. Seventeen abstractors accounted for nearly half (112) of the hospitals. The greatest number of hospitals abstracted by a single abstractor was nine.

d. On-Site Data Collection

Data collection for the study began August 21, 1982 and continued through November 30 (with one final hospital completed on December 26). Data collection was initiated by contacting each hospital's medical records department to establish a date for abstracting. In order to minimize abstractor training and travel costs, all hospitals in a given area were generally contacted at the same time.

At this time, we also asked that the medical records department supply us with the information necessary for developing the sampling plan for their hospital. This included information on the total discharges in 1980, the number of normal deliveries and pediatric cases, and the type of discharge records available for sampling. A sampling plan for the hospital was then computer-generated by project staff and mailed along with detailed instructions to the hospital at least a week in advance of their abstracting date. (See Appendix G for examples of both types of Sampling Plans.)

Concurrent with the scheduling of hospitals, abstractors were contacted and trained. Once an abstractor was assigned a hospital, he/she was mailed a packet containing the materials necessary to complete the assignment. For identification purposes, each abstractor received a letter of introduction from SysteMetrics to the Director of Medical Records, and a copy of the Agreement of Participation signed by the hospital's administrator. Additionally, each abstractor was asked to sign a confidentiality agreement and return it. A copy of this agreement was then sent to the appropriate hospital along with the sampling plan to assure the hospital that confidentiality measures were adequate. The "field packet" also included an abstracting record and a set of pre-numbered abstracting booklets (SMI's) to be used in the assigned hospital only. A log was kept of which pre-numbered SMI's were issued to each hospital to aid in subsequent hospital identification. In addition, each SMI carried a two-digit number identifying the abstractor or "rater" and a three-digit number identifying the hospital.

The abstractor's first task upon arrival at the hospital was to reorder the charts pulled by the medical records staff in a random fashion. Each abstractor was given computer-generated instructions on how to do this (shown in Appendix H). The abstractor then abstracted each record in the new random order given on the abstracting record.

Before abstracting a record, the rater determined whether it should be retained in the sample or not. The following types of records were discarded:

- normal deliveries (i.e., deliveries with no complications, which had the ICD-9-CM diagnostic code of 650)
- pediatric cases (i.e., those patients discharged from the pediatric service or whose age was 15 or less)
- records for patients not discharged in 1981
- records for patients not occupying an acute care bed
- incomplete records (i.e., those with no discharge summary in the chart)
- records removed from the stack by the Medical Records Department

Information was abstracted from each medical record in the order given by the various sections of the Standardized Medreview Instrument. The first component was the Patient Background section, consisting of routine information describing the patient and an overview of the admission, treatment milestones and discharge. Information was collected on:

- admit date
- admit time
- discharge date
- discharge time
- date of birth
- expected source of principal payment
- primary placement (e.g., medical, surgical, ICU, etc.)
- source of admit (e.g., routine, emergency room)
- disposition of patient
- diagnoses
- procedures and dates performed

Abstractors were instructed to make reasonable inferences regarding the source of admission and the disposition of the patient if these were unavailable from the face sheet. In addition, if the date of birth were unavailable, abstractors were asked to approximate patient's age and to indicate that this had been done.

Abstractors were asked to provide up to three diagnoses: the principal diagnosis, and two of the patient's other major diagnoses pertinent to the admission being abstracted. UHDDS guidelines were used in designating the principal and secondary diagnoses. The UHDDS definition of principal diagnosis is: "The condition established after study to be chiefly responsible for occasioning the admission of the patient to the hospital for care." Secondary diagnoses are "all conditions that exist at the time of admission or develop subsequently which affect the treatment received and/or length of stay. Diagnoses that relate to an earlier episode which have no bearing on this hospital stay are to be excluded."

Abstractors were also instructed to record up to 4 major (i.e., Class 1) procedures performed on a patient, with the principal procedure listed first. The question as to which procedure to list as the principal one was also determined by the UHDDS definition: When more than one procedure is recorded, the principal procedure is determined by applying the following criteria:

1. The principal procedure is one which was performed for definitive treatment rather than one performed for diagnostic or exploratory purposes, or was necessary to take care of a complication.
2. The principal procedure is that procedure most related to the principal diagnosis.

After completing the Patient Background information, the abstractor then turned to the remaining sections of the instrument including the Admission Review, the Level of Service Review, and the Need for Continued Hospitalization. Any special considerations were discussed in detail in the Notes section of the instrument. This section was provided for the abstractor to write out as much detail as needed to explain or justify any override decision or to provide any other information necessary for evaluating the case.

The number of records abstracted by each rater varied from hospital to hospital. A maximum number of 25 abstracts were considered for data analysis, and the minimum number judged to constitute a reliable sample was 15. Abstractors were instructed to continue abstracting until they had gone through all the records on their list or until they had spent two days at the hospital. Most abstractors had no trouble completing their assignment within the time allowed.

Upon completing the assignment, each abstractor was instructed to return to SysteMetrics the abstracting record, the sampling plan (as completed by the hospital), an invoice for time and travel expenses, and all the SMI abstracting booklets, completed or not. Unused abstracts were not reassigned to other hospitals.

e. Quality Control Procedures

All abstractors were instructed to call SysteMetrics if any problems or questions arose while abstracting in the hospitals and were allowed to call collect. Additionally, abstractors were instructed to use the Notes section to provide details of cases for which they had questions. These questions could then be answered by project staff or referred to the consultant physician for clarification. One example of the type of question that arose was "If a patient is admitted for gastroenteritis with dehydration but does not have any abnormal (according to our criteria) electrolyte values, is the admission justified?" This was referred to one of the consultant physicians who indicated that the admission should be judged as justified.

Periodic memoranda were sent to all abstractors reporting this and other necessary clarifications, as well as noting how to solve problems that might be encountered such as, how to use the abstracting record if more than 50 medical records had been pulled for them. Further clarifications on use of the instrument were also included.

Once the data collected in each hospital was returned to SysteMetrics, each abstract was reviewed for technical consistency and any abstractor notes or questions were incorporated into the data. This review consisted of an examination of the diagnoses, procedures and criteria checked to ensure that all were consistent. If, for instance, the admission criterion for cardiac pathology was checked, one of the diagnoses listed should have indicated cardiac disease. If an inconsistency was found, the abstractor was called for clarification. This occurred infrequently. If the abstractor was unable to provide the information, the hospital was then called for clarification. This happened only twice. Occasionally, abstractors neglected to complete the date of birth of the patient. If the age for the patient could not be approximated from the notes, the hospital was called. This happened more frequently -- in a total of about ten hospitals. Technical review of the charts also involved checking Class I procedure dates in three places for consistency, and making any written comments for data entry that might be pertinent to a case. All abstracts were then coded to correspond with the protocol written for key punching the data. After key entry, edit checks were performed on the data. These included the following:

- Admission date should be before or the same as the discharge date.
- Procedure dates and all days abstracted should fall within the period between admission and discharge dates.
- Patient Age should not be less than 16 years (the definition of a pediatric case for this study).
- Days abstracted for level of service and for continued hospitalization should be identical.
- If certain admission criteria were checked, diagnostic codes must be within specified ranges, (e.g., cardiac pathology, status asthmaticus, fractures).
- Admission and discharge times could only be within specified ranges (1-24).

- the combination of criteria checked or left blank and subsequent appropriateness evaluations checked were compared and had to be consistent.
- If an admission or continued hospitalization day was evaluated as not justified, or a level of service on any day was evaluated as non-acute, then at least one reason code within specified ranges must be present.
- Identification data of Sequential Number, Rater Number and Hospital Number must occur in specified combinations.

About ten percent of the data failed these edit checks, primarily on the basis of key entry errors or incomplete dates. These were corrected on the data files, following an examination of the abstract to ascertain the correct values.

IV. DATA ANALYSIS AND METHODOLOGICAL ISSUES

IV. DATA ANALYSIS AND METHODOLOGICAL ISSUES

A. DESCRIPTION OF THE SAMPLE

The final data set was comprised of 5,732 admissions and 24,051 days of care, sampled from 253 hospitals for patients discharged in 1981. The following section discusses the characteristics of sampled hospitals and patient records. This is followed by sections evaluating response bias and sample representativeness.

1. Sample Characteristics

Exhibit IV-1 presents the unweighted and weighted distribution of sampled hospitals by region, bedsize, ownership, SMSA-size, teaching status, and occupancy rate. Also shown is the distributions of hospitals in the population. As noted in the previous chapter, hospitals participating in the study represented all geographic regions within the continental United States. Only non-federal general short-term hospitals were included in the sample.

Hospitals were sampled with probability proportionate to their 1980 admissions so the weight for each hospital is the stratum "measure of size" (i.e., the total number of 1980 admissions for hospitals included in the stratum) divided by the number of 1980 admissions for the sampled hospital. Because hospitals were sampled with probability proportional to size, any hospital characteristic positively associated with size will be over-represented in the unweighted sample, but will not be disproportionately represented when the sample is properly weighted. Exhibit IV-1 shows, in fact, that the distribution of hospitals, when weighted, is very similar to that in the population across all the characteristics examined.

With respect to the unweighted sample, the greatest number of hospitals (39%) were sampled from the Southern region, followed by the North Central region (25%), Northeastern region (19%), and Western region (17%). Thirty percent of the hospitals (unweighted) had over 400 beds while only 4 percent had

EXHIBIT IV-1

Unweighted Distribution of Sampled Hospitals, Weighted Distribution of Sampled Hospitals, and Distribution of Hospitals in the Population by Selected Hospital and Area Characteristics

	Unweighted Distribution of Sampled Hospitals (n=253)	Weighted Distribution of Sampled Hospitals	Population Distribution (N=5696)
REGION	Percent		
Northeast	18.6%	16.3%	14.6%
North Central	25.3%	24.0%	29.6%
South	39.1%	39.5%	37.6%
West	17.0%	20.3%	18.2%
	100.0%	100.0%	100.0%
BEDSIZE			
0 - 49	4.3%	17.5%	22.3%
50 - 149	23.7%	43.2%	39.0%
150 - 249	18.6%	14.6%	16.2%
250 - 399	23.3%	13.2%	12.2%
400 +	30.3%	11.6%	10.2%
	100.0%	100.0%	100.0%
OWNERSHIP			
Government (nonfederal)	19.0%	34.4%	31.4%
Proprietary	18.6%	13.0%	12.9%
Nonprofit	62.5%	52.6%	56.3%
	100.0%	100.0%	100.0%
SMSA SIZE			
Non SMSA	24.9%	48.4%	51.1%
< 100K	2.4%	1.4%	1.7%
100 - 250K	7.1%	4.7%	7.1%
250 - 500K	12.6%	8.4%	8.6%
500K - 1mil	12.3%	8.0%	7.2%
1 - 2.5mil	24.5%	15.1%	13.2%
> 2.5mil	16.2%	13.9%	11.6%
	100.0%	100.0%	100.0%
TEACHING STATUS			
Teaching	36.0%	16.3%	16.9%
Nonteaching	64.0%	83.7%	83.1%
	100.0%	100.0%	100.0%
OCCUPANCY RATE			
< 39%	2.0%	11.3%	6.7%
40 - 59%	9.9%	23.6%	22.8%
60 - 79%	43.1%	38.3%	43.1%
80 - 89%	34.4%	21.7%	21.3%
90% +	10.7%	5.1%	6.1%
	100.1%	100.0%	100.0%

n = number of hospitals in sample

N = number of hospitals in population

fewer than 50 beds. When weighted, large hospitals (over 400 beds) comprise 12% of the sample, while small hospitals (under 50 beds) comprise 18% of the sample.

In terms of ownership, the largest category represented in the sample are nonprofit hospitals which comprise 62% of the total. As discussed previously, proprietary hospitals were over-sampled, and so comprise 19% of the unweighted distribution compared to 13% in the population. Hospitals from non-SMSA areas represented nearly a quarter of the unweighted sample, while those from areas with over a million in population comprise almost 41% of the sample. Non-teaching hospitals (defined as those without a medical school affiliation or a residency program) are the largest category (64%) in the sample. Because of the sampling strategy, however, teaching hospitals at 36% are more numerous in the unweighted sample than in the weighted sample or the population (in which only 17% of hospitals can be defined as teaching hospitals). Finally, with respect to occupancy rate, most sampled hospitals fell in the mid-range. Over three-fourths of the hospitals had an occupancy rate between 60% and 89%, whereas approximately 12% of the sampled hospitals had occupancy rates less than 60% and another 10% had occupancy rates greater than 89%.

Exhibit IV-2 presents the unweighted distribution of sampled patient records by selected patient and hospital stay characteristics, including patient age, sex, paysource, source of admission, admission placement, patient disposition, month of discharge, month of admission, length of stay, performance or nonperformance of a major procedure during the stay and the major ICD-9-CM class of the primary diagnosis. (Weighted distributions will be presented later in the chapter.) In addition, any patient experiencing a normal delivery, discharged from a pediatric service, or who was less than 16 years of age was excluded from the sample. All other cases had an equal chance of being selected.

As can be seen in Exhibit IV-2 the sample had a fairly even distribution across age categories, with the heaviest concentration (32%) being for those 65 years or older (16% of the sample was 75 years or older). Twenty-four observations in the sample had age missing through abstracting error. Females outnumbered males in the sample by 17.5%, and 3 abstract records lacked information as to the patient sex. Medicare (33%), Blue Cross (18%), and

EXHIBIT IV-2
Unweighted Distribution of Sampled Patients
by Selected Patient and Hospital Stay Characteristics

PATIENT AGE	Frequency	Percent
16 - 29	1280	22.3%
30 - 44	1115	19.5%
45 - 64	1520	26.5%
65 - 74	899	15.7%
75 +	894	15.8%
Not Available	24	0.0%
	5732	100.0%
SEX		
Male	2362	41.2%
Female	3367	58.7%
Not Available	3	0.0%
	5732	100.0%
PAYSOURCE		
Medicare	1897	33.1%
Medicaid	393	6.9%
Blue Cross	1014	17.7%
Commercial	1567	27.3%
Other Government	99	1.7%
Workman's Compensation	112	2.0%
Self Pay	234	4.1%
HMO	41	0.7%
Other	120	2.1%
Not Available	255	4.4%
	5732	100.0%
SOURCE OF ADMISSION		
Routine	3492	60.9%
Emergency Room	1920	33.5%
Skilled Nursing Facility	93	1.6%
Other Acute Facility	84	1.5%
Ambulatory Surgery	15	.3%
Not Available	128	2.2%
	5732	100.0%
ADMISSION PLACEMENT		
Medical	2558	44.6%
Surgical	2629	45.9%
Psychiatric	104	1.8%
ICU or CCU	346	6.0%
Burn Care Unit	7	0.1%
Alcoholism/Chemical Dependency	27	0.5%
Organized Rehabilitation Unit	7	0.1%
Trauma/Shock Unit	5	0.1%
Other	19	0.3%
Not Available	30	0.5%
	5732	100.0%

EXHIBIT IV-2 cont.

DISPOSITION OF PATIENT	FREQUENCY	PERCENT
Routine	4959	86.5%
Skilled Nursing Facility	191	3.3%
Died	178	3.1%
Acute Facility	112	2.0%
Non-Acute Facility	99	1.7%
Home Care	74	1.3%
AMA	55	1.0%
Not Available	64	1.1%
	5732	100.0%
MONTH AND YEAR OF DISCHARGE		
1981:		
Jan	494	8.6%
Feb	444	7.8%
March	489	8.5%
April	508	8.9%
May	518	9.0%
June	512	8.9%
July	483	8.4%
August	479	8.4%
September	483	8.4%
October	484	8.4%
November	436	7.6%
December	400	7.0%
1982:		
Jan	1	.0%
	5732	99.9%
MONTH AND YEAR OF ADMISSION		
1980:		
June	1	.0%
October	1	.0%
November	4	.1%
December	81	1.4%
1981:		
Jan	509	8.9%
Feb	468	8.2%
March	493	8.6%
April	513	9.0%
May	507	8.8%
June	513	9.0%
July	469	8.2%
August	471	8.2%
September	484	8.4%
October	466	8.1%
November	443	7.7%
December	308	5.4%
	5732	100.0%

EXHIBIT IV-2 cont.

LENGTH OF STAY (in days)	FREQUENCY	PERCENT
1	553	9.65%
2	820	14.29%
3	721	12.29%
4	607	10.59%
5	494	8.62%
6	409	7.14%
7	351	6.12%
8 - 10	697	12.16%
11 - 20	761	13.28%
21 - 30	185	3.23%
31 - 99	127	2.22%
100 +	7	.12%
	5732	100.00%
MAJOR PROCEDURE		
Performed	2794	48.8%
Not Performed	2938	51.2%
	5732	100.0%
MAJOR ICDA CLASS		
Infectious & Parasitic Diseases	78	1.4%
Neoplasms	403	7.0%
Endocrine, Nutritional & Metabolic Dis.	155	2.7%
Diseases of the Blood	36	.6%
Mental Disorders	224	3.9%
Disease of the Nervous System	272	4.8%
Diseases of the Circulatory System	876	15.3%
Diseases of the Respiratory System	449	7.8%
Diseases of the Digestive System	746	13.0%
Diseases of the Genitourinary System	631	11.0%
Complications of Pregnancy, Childbirth	444	7.7%
Diseases of the Skin	84	1.5%
Diseases of the Musculoskeletal System	372	6.5%
Congenital Anomalies	20	.4%
Symptoms, Signs & Ill-defined Conditions	309	5.4%
Injury and Poisoning	534	9.3%
Other	99	1.7%
	5732	100.0%

Commercial Insurance (27%) were by far the largest payors in the sample. Only 41 cases (0.7%) in the sample represented an HMO pay source. "Other government," "self-pay," and "other" were also small categories representing 1.7%, 4.1%, and 2.1% of the total respectively. It should be noted that pay source was the patient background variable with the highest missing data rate (4.4%).

With respect to admission source, admissions from other facilities were rare in the sample. (A total of 3.4% of the sampled patients were admitted from skilled nursing facilities, other acute facilities, or ambulatory surgery units.) The great majority of sampled admissions were routine (61%) or through the emergency room (34%). For 2.2% of the admissions, there was no available indication of admission source in the medical record.

Most sampled admissions were placed either in a medical (45%) or surgical service (46%). Placement in an ICU or CCU unit (6%) or in a psychiatric unit (2%) comprised the next largest categories. Only 31 cases total (0.5%) fell into the remaining categories (burn care unit, alcoholism dependency unit, organized rehabilitation unit, trauma/shock unit, other unit). For 30 cases (0.5%), admission placement was not available.

Most patient discharges in the sample were routine (87%). Only a combined total of 364 (6%) were to lower levels of care (skilled nursing facility, other non-acute facility, or home care) while 112 (2%) were discharged to other acute facilities. Patients who died or who left against medical advice were also small categories in the sample (3% and 1% respectively). In slightly over 1% of the records patient disposition could not be determined.

The date of patient discharge was approximately equally distributed over each month of the year, with a slight decrement in November and December. (One patient in the sample was discharged on January 1, 1982). Admissions were generally in 1981, although 87 patients (1.5%) were admitted in 1980 and discharged in 1981. Almost all of these were admitted in December of 1980. The majority of sampled stays (81%) were ten days or less. Only 134 (2%) were over 30 days and only 7 (0.1%) were over 100 days. The longest stay in the sample was 343 days.

In almost half (49%) of the stays major surgery (defined as a UHDDS class 1 procedure) was performed. As shown in Table IV-2, the most frequent diagnostic classes represented in the data were Diseases of the Circulatory System (15%), Diseases of the Digestive System (13%), Diseases of the Genitourinary System (11%) and Injuries and Poisonings (9%).

The most frequently occurring diagnosis (three-digit ICD-9-CM) in the sample was ICD-9-CM code 414 (other forms of chronic ischemic heart disease) with 127 or (2.2%) of the cases. This was followed by ICD-9-CM codes 366 (cataract) with 1.7% of the cases, 410 (acute myocardial infarction) with 1.6% of the cases, 574 (cholelithiasis) with 1.6% of the cases, 428 (heart failure) with 1.5%, 786 (symptoms involving the respiratory system and chest) with 1.5% of the cases, 550 (inguinal hernia) with 1.4% of the cases, 427 (cardiac dysrhythmias) with 1.4% of the cases, and 626 (disorders of menstruation and other abnormal bleeding from the female genital tract) with 1.4% of the cases.

2. Response Bias

Exhibit IV-3 presents the response distribution of hospitals. Those that responded to the initial mailout request are labelled the primary sample and those that responded to subsequent mailout requests are labelled the shadow samples. As can be seen, 164 of the 250 non-HMO study hospitals were respondents to the original participation request, for a response rate of 66%. Of the 15 hospitals drawn from the HMO sample, only 3 participated, for a response rate of 20%. Since all 12 HMO nonrespondents were members of the same organization, no follow-up hospitals were selected to replace the nonresponders in this sample.

Follow-up hospitals were recruited as substitutes for hospitals in the initial mailing who declined to participate. Each hospital was chosen from the same stratum as the original hospital and so was equivalent in terms of region and ownership. In this process, follow-up hospitals were selected to match the original hospitals as closely as possible in terms of bedsize, teaching status,

EXHIBIT IV-3

DISTRIBUTION OF HOSPITALS FROM PRIMARY AND SHADOW SAMPLES

	Frequency	Percent
Non-HMO Sample		
Primary Sample (i.e., responded to original mailout)	164	65.6%
First Shadow Sample	58	23.2%
Second Shadow Sample	20	8.0%
Third Shadow Sample	5	2.0%
Fourth Shadow Sample	3	1.2%
Total	250	100.0%
HMO Sample		
Primary Sample Respondents	3	20.0%
Non-respondents	12	80.0%
Total	15	100.0%

SMSA-size and occupancy rate. Because of this the final sample distribution will necessarily resemble the distribution of the primary sample on bedsize, region, ownership, and the other stratification variables, and so retain the preferred distribution on these dimensions.

Although the final sample will retain the national representativeness that was originally built into the sampling strategy, nonresponding hospitals may nevertheless represent a different subpopulation than responding hospitals. Exhibit IV-4 provides a comparison of the primary responding and primary nonresponding hospitals in terms of bedsize, occupancy rate, region, ownership, teaching status, and SMSA size. Primary responding hospitals are defined as those that agreed to participate following the initial mailing. Primary nonresponding hospitals are those that declined to participate at that time. T-tests were used to assess mean differences between the two samples with respect to continuous variables such as bedsize; Chi-square tests were used for categorical variables such as region.

Although there were some apparent differences between the two samples, they nevertheless appeared similar on most dimensions. As can be seen in Exhibit IV-4, there were no significant differences between responding and nonresponding hospitals in terms of bedsize or any of the utilization measures, including admission rate, days of care rate, occupancy rate, and average length of stay. Primary responders, in addition, were not drawn differentially from any of the four regions. They were, however, from the medium-sized SMSAs (100,000 to 1,000,000) significantly more often ($p < .05$) than the nonresponders: 76% of the hospitals from medium-sized SMSAs were primary responders compared with 62% from small and 59% from large SMSAs. There was also a tendency for primary responders to be teaching rather than nonteaching hospitals and to be government-owned or nonprofit rather than profit hospitals. Neither of these relationships, however, achieved significance.

EXHIBIT IV-4

Comparison of Responding and Nonresponding Sampled Hospitals by Selected Hospital and Area Characteristics

	Responding Hospitals n = 167	Nonresponding Hospitals n = 86	
	mean		t
BEDSIZE	338.4	319.6	0.59^{ns}
1980 ADMISSIONS	12662.2	12009.3	0.53^{ns}
1980 DAYS OF CARE	98203.4	91717.3	0.60^{ns}
OCCUPANCY RATE	75.9	24.3	0.94^{ns}
AVERAGE LOS	7.5	7.2	1.34^{ns}
REGION	percentage		X^2
North Central	65.6%	34.4%	2.08^{ns}
Northeast	68.1%	31.9%	
South	61.6%	38.4%	
West	73.8%	26.2%	
OWNERSHIP			
Government	66.7%	33.3%	4.84^{ns}
Non-profit	69.6%	30.4%	
Proprietary	52.2%	47.8%	
TEACHING STATUS			
Teaching	73.6%	26.4%	3.81^{ns}
Non-teaching	61.5%	38.5%	
SMSA SIZE			
< 100K	62.0%	38.0%	6.44*
100K - 1mil	75.8%	24.2%	
> 1mil	58.9%	41.1%	

ns = non significant

* = significant at the .05 level

3. Representativeness of the Sample

Two sources are available for assessing the representativeness of the national sample drawn for the study. The first is the National Hospital Discharge Survey (NHDS), conducted annually since 1964. The second is the American Hospital Association's (AHA) 1981 Annual Survey of Hospitals which covers all hospitals in the nation. Data tapes for these two surveys were acquired for the year 1981 and population estimates for these samples compared with population estimates derived from this study's sample (termed the SMI sample).

Three quantities were estimated: (1) total admissions, (2) total days of care, and (3) average length of stay. These were examined by region, ownership category, and bedsize. The SMI sample was weighted using the sampling weights based on the original probabilities of selection and adjustments for nonresponse described in the following section. The first- and second-stage ratio estimators were not applied for this original analysis of the representativeness of the sample since these estimates would have automatically produced distributions by region, ownership, and bedsize category identical to those of the NHDS sample.

The National Hospital Discharge Survey was weighted using the variable provided on the tape record for that purpose. The Annual Hospital Survey is a complete census of hospitals, each of which reports its annual admissions and days of care. To obtain aggregate admissions and days of care by each of the region, ownership, and bedsize categories for the AHA survey, these quantities were summed across all institutions in a given category.

As noted, the SMI sample excludes all specialty hospitals as well as all pediatric patients and patients with normal deliveries. The National Hospital Discharge Survey includes short term specialty as well as general hospitals. Because there is no identifier for such hospitals on the data tape for the survey, these could not be excluded from the NHDS data for the present analysis.

Pediatric patients, however, were excluded on the basis of age (age < 16 years) and normal deliveries were excluded on the basis of primary diagnosis code (ICD-9-CM code 650). When tabulating admissions and days of care from the AHA Annual Survey of Hospitals, specialty hospitals could be identified and excluded on the basis of the "Type of Service" code. There was no reliable indicator of the number of pediatric discharges, however, or the number of normal deliveries. In addition, AHA statistics may also include utilization of the subacute unit of a facility. These admissions were not always reliably indicated on the AHA Survey. Both the NHDS and the AHA survey, therefore, cover larger populations than those covered by the SMI sample, so that the admissions and days of care estimated in these surveys will be higher automatically.

Exhibit IV-5 presents the results of this analysis. Shown first are admissions, days of care and average length of stay for the SMI sample, followed by those taken from the AHA and NHDS. Overall, the SMI sample estimates 28,760,221 discharges and 218,795,910 days of care for adult patients discharged from general hospitals in 1981 for reasons other than normal delivery. The AHA estimated 35,714,899 admissions and 272,677,170 days of care for all patients discharged from general hospitals in 1981, a figure which may include admissions and days of care for the subacute unit of the facility. Finally, the NHDS estimates 34,532,257 admissions and 258,531,270 days of care in specialty and general hospitals in 1981 for adult patients admitted for reasons other than normal deliveries. (Subsequent analysis showed that normal deliveries were not necessarily coded as ICD-9-CM code 650 in the NHDS data set and so may be represented in the statistics calculated above.)

The distribution of admissions by region is remarkably similar for all three surveys. Approximately 28% of the admissions in each sample are in the North Central region, 20% in the Northeast, 35% in the South and 16% in the West. For days of care, these numbers are 30%, 24%, 32%, and 14% respectively. The average length of stay (ALOS) by region is also similar across the three samples, and mirrors the same trend noted by previous investigators, namely that ALOS is greatest in the Northeastern region (almost 9 days) and lowest in the Western region (slightly over 6 days). The North Central region has the next

EXHIBIT IV-5

Comparison of Estimated Admissions, Days of Care and ALOS with the 1981 American Hospital Association Annual Survey of Hospitals (AHA) and 1981 National Hospital Discharge Survey (NHDS)

	SMI SURVEY					AHA					NHDS				
	Admissions	%	Days of Care	%	ALOS	Admissions	%	Days of Care	%	ALOS	Admissions	%	Days of Care	%	ALOS
REGION															
North Central	8,076,208	28.1%	64,411,767	29.4%	8.0	9,919,776	27.8%	80,052,519	29.4%	8.1	9,940,132	28.8	76,855,069	29.7%	7.7
Northeast	5,836,230	20.3%	51,184,646	23.4%	8.8	7,241,840	20.3%	64,474,105	23.6%	8.9	7,077,321	20.5	62,655,724	24.2%	8.9
South	10,482,889	36.4%	76,247,419	34.8%	7.3	12,860,497	36.0%	91,475,647	33.5%	7.1	11,675,233	33.8	83,031,792	32.1%	7.1
West	4,364,884	15.2%	26,952,092	12.3%	6.2	5,692,786	15.9%	36,674,896	13.4%	6.4	5,839,571	16.9	35,988,688	13.9%	6.2
OWNERSHIP															
Proprietary	2,750,977	9.6%	18,461,315	8.4%	6.7	3,208,677	9.0%	21,006,274	7.7%	6.5	2,827,236	8.2%	21,596,429	8.4%	7.6
Nonprofit	20,184,270	70.2%	158,458,831	72.4%	7.9	25,301,761	70.8%	197,816,122	72.5%	7.8	24,359,992	70.5%	186,898,210	72.3%	7.7
Government (non-federal)	5,824,964	20.2%	41,875,779	19.1%	7.2	7,204,461	20.2%	53,854,771	19.8%	7.5	7,345,029	21.3%	50,036,634	19.4%	6.8
BEDSIZE															
0-99	4,263,487	14.8%	26,426,464	12.1%	6.2	4,839,047	13.5%	31,475,739	11.5%	6.5	6,281,839	18.2%	38,672,861	15.0%	6.2
100-299	9,858,769	34.3%	72,190,377	33.0%	7.3	13,377,204	37.5%	97,374,163	35.7%	7.3	11,572,074	33.5%	83,742,840	32.4%	7.2
300-499	8,332,836	29.0%	63,783,973	29.2%	7.7	9,735,246	27.3%	75,945,482	27.9%	7.8	8,534,081	24.7%	68,344,817	26.4%	8.0
500+	6,305,118	21.9%	56,395,110	25.8%	8.9	7,763,402	21.7%	67,881,783	24.9%	8.7	8,144,263	23.6%	67,770,755	26.2%	8.3
TOTAL	28,760,221	100.0%	218,795,910	100.0%	7.6	35,714,899	100.0%	272,677,170	100.0%	7.6	34,532,257	100.0%	258,531,270	100.0%	7.5

[1] Excludes specialty hospitals and pediatric and normal delivery admissions.

[2] Excludes specialty hospitals, includes pediatric and normal delivery admissions, and some admissions to the subacute unit.

[3] Includes specialty hospitals, excludes pediatric and normal delivery admissions.

highest ALOS (at around 8 days), while the South has the next lowest (at slightly over 7 days).

Exhibit IV-5 also shows very similar distributions across the three samples by ownership category and bedsize for both admissions and days of care. Average length of stay fluctuates somewhat more for the ownership and bedsize categories than for region, but the figures are generally quite consistent. In addition, the three estimates of ALOS for the nation as a whole are within a tenth of a percent of each other at 7.6 days for the SMI and AHA surveys and 7.5 days for NHDS.

Exhibit IV-6 provides a further comparison of the SMI sample with the NHDS survey for a number of patient and hospital stay characteristics. Presented are estimated admissions by patient age, sex, principal pay source, patient disposition, length of stay, the performance or nonperformance of a major procedure, and major ICDA class. As can be seen, the distribution by age and sex in both samples is quite similar. With respect to age, there is an approximately even distribution across age classes, with the greatest percentage falling in the 65 years or older category (32% in the SMI sample, 30% in the NHDS). Females predominate in both samples with 59% of the SMI sample and 62% of the NHDS being female. In both samples Medicare is the major payor (34% for the SMI and 33% for NHDS) followed by commercial insurance (27% and 28% respectively) and Blue Cross (19% and 21%). Medicaid accounts for 7% of the SMI survey and 8% of the NHDS. The remaining pay categories are also fairly equivalent.

Both samples show only a small percentage of patients being discharged to long term care facilities (3% in each sample), with the great majority being discharged routinely (87% of the SMI survey and 82% of the NHDS). The NHDS, however, had a much higher percentage of patients for whom discharge status could not be determined (10%) than did the SMI (1%).

The distribution by length of stay is also very similar for the two samples. One day stays accounted for around 10% of all admissions in the NHDS

EXHIBIT IV-6

Comparison of Weighted Admissions by Selected Patient Characteristics with the 1981 National Hospital Discharge Survey (NHDS)

	SMI SURVEY		NHDS SURVEY	
PATIENT AGE	Frequency	Percent	Frequency	Percent
16-29	6,290,365	22.0%	8,935,378	25.9%
30-44	5,456,950	19.1%	6,511,107	18.9%
45-64	7,747,869	27.1%	8,677,432	25.1%
65+	9,134,423	32.0%	10,408,340	30.1%
	28,629,607	100.2%	34,532,257	100.0%
SEX				
Male	11,854,386	41.2%	13,158,582	38.1%
Female	16,900,592	58.7%	21,373,675	61.9%
Not Available	12,731	.0%	-	.0%
	28,767,709	99.9%	34,532,257	100.0%
PAYSOURCE				
Medicare	9,670,558	33.6%	11,327,460	32.5%
Medicaid	1,904,709	6.6%	2,698,675	7.8%
Blue Cross	5,326,749	18.5%	7,354,601	21.3%
Commercial	7,615,056	26.5%	9,812,122	28.4%
Other Government	455,189	1.6%	645,759	1.9%
Worker's Compensation	560,360	2.0%	641,938	1.9%
Self Pay	1,160,898	4.0%	1,705,333	4.9%
HMO	176,281	.6%	N/A	N/A
Other	581,962	2.0%	436,369	1.3%
Not Available	1,316,190	4.6%	N/A	N/A
	28,767,952	100.0%	34,532,257	100.0%
DISPOSITION OF PATIENT				
Routine	25,076,942[1]	87.2%	28,360,644	82.1%
Long Term Care	951,087[2]	3.3%	1,017,293	2.9%
Died	963,704	3.4%	949,078	2.7%
Acute Facility	522,129	1.8%	489,820	1.4%
AMA	890,661	3.1%	288,615	.8%
Not Available	351,353	1.2%	3,426,807	9.9%
	28,755,876	100.0%	34,532,257	99.8%

[1] To create a category equivalent to the NHDS classes, "Routine" and "Home Care" are combined.

[2] SNF and other non-acute facility are combined.

Exhibit IV-6 cont.

	SMI SURVEY		NHDS SURVEY	
LENGTH OF STAY (in days)	Frequency	Percent	Frequency	Percent
1	2,712,171	9.4%	3,531,749	10.2%
2	4,027,445	14.0%	4,835,112	14.0%
3	3,636,691	12.6%	4,674,915	13.5%
4	3,088,410	10.7%	3,667,318	10.6%
5	2,468,022	8.6%	2,807,097	8.1%
6	1,991,444	6.9%	2,296,457	6.6%
7	1,771,561	6.2%	1,955,395	5.7%
8-10	3,487,309	12.1%	3,924,989	11.4%
11-30	4,872,543	16.9%	5,984,737	17.3%
31+	712,351	2.4%	854,488	2.5%
	28,629,607	100.0%	34,532,257	99.9%
MAJOR PROCEDURE				
Performed	28,767,947	48.8%	15,020,286	43.5%
Not Performed	14,730,287	51.2%	19,511,971	56.5%
	28,629,607	100.0%	34,532,257	100.0%
MAJOR ICDA CLASS				
Infectious and Parasitic Dieases	397,018	1.4%	454,374	1.3%
Neoplasms	2,023,273	7.0%	2,509,694	7.3%
Endocrine, Nutritional & Metabolic Dis.	774,707	2.7%	1,085,475	3.1%
Diseases of the Blood	182,230	.6%	293,279	.8%
Mental Disorders	1,147,118	4.0%	1,679,334	4.9%
Disease of the Nervous System	1,394,056	4.8%	1,487,833	4.3%
Diseases of the Circulatory System	4,491,227	15.6%	5,289,418	15.3%
Diseases of the Respiratory System	2,231,069	7.8%	2,402,179	7.0%
Diseases of the Digestive System	3,830,754	13.3%	4,139,952	12.0%
Diseases of the Genitourinary System	3,071,373	10.7%	3,310,144	9.6%
Complications of Pregnancy, Childbirth	2,164,279	7.5%	991,212	2.9%
Diseases of the Skin	405,959	1.4%	506,722	1.5%
Diseases of the Musculoskeletal System	839,691	6.4%	2,204,980	6.4%
Congenital Anomalies	119,109	.4%	146,030	.4%
Symptoms, Signs & Ill-defined Conditions	1,496,224	5.2%	522,025	1.5%
Injury and Poisoning	2,705,450	9.4%	3,039,808	8.8%
Supplementary Classifications (V Codes)	494,410	1.7%	4,469,798	12.9%
	28,767,947	99.9%	34,532,257	100.0%

and 9% in the SMI survey, while stays over 30 days were around 2.5% in both. The NHDS indicated a somewhat lower rate of major surgery showing only 44% of the admissions having a class 1 procedure (as opposed to 49% in the SMI survey). The distribution of primary diagnoses across the major ICDA classes was also quite similar for the two surveys with three exceptions: The NHDS showed a lower rate for complications of pregnancy (3% versus 8%) and a lower rate for symptoms and ill-defined conditions (2% versus 5%) but a much higher rate for the supplementary classifications (i.e., the ICD-9-CM V codes). This is most likely due to differences in the coding practices of the two surveys, in that a large number of pregnancies appear to have been coded in the NHDS with the code V27 (Outcome of Delivery). In addition, approximately 1% of the NHDS codes were V71, Observation and Evaluation for Suspected Conditions, including that for suspected cardiovascular disease. These are cases that would otherwise be coded under the ICD class, symptoms and ill-defined conditions. This may account for the lower percentage of symptoms and ill-defined conditions in the NHDS sample than the SMI sample.

B. ESTIMATION AND CALCULATION OF SAMPLING VARIANCE

The national sample, being a complex one, required a sophisticated procedure for deriving estimates and calculating the sampling precision of these estimates. The estimation procedure involved first the calculation of the initial sampling weights and non-response adjustments. Following this the estimation procedure was further refined by the calculation of first and second stage ratio estimators using the National Hospital Discharge Survey and the 1981 AHA Annual Survey of Hospitals to provide independent estimates of days of care by division, ownership category, bedsize, and medical/surgical status.

Given the complex multi-stage cluster sampling approach used to select hospitals, no direct computational procedure for obtaining sampling variance could be applied feasibly. The study, therefore, used a replication method of estimating the sampling variance of a statistic from multiple subsamples within the parent sample. This was followed by development of a generalization function so that sampling error could be estimated across a wide range of values without having to apply the replication algorithm for each statistic in the analysis. Each of these steps will be discussed in turn. Appendix I provides practical examples of the calculations and procedures described below.

1. Estimation Procedures

a. Calculation of Initial Sampling and Nonresponse Weights

Each admission and day of care in the sample was assigned a weight representing its contribution to the population total. The overall weight was first a product of the inverse of the initial sampling probabilities, which were (1) the probability of selecting the hospital and (2) the probability of selecting a record in the hospital. Three subsequent adjustments for nonresponse and selection came into play once the medical records identified for inclusion in the study were pulled from the shelf. These five weights will be described below.

A given hospital was selected with probability proportional to size within a given stratum, so that the probability of selection was the hospital measure of size (i.e., 1980 total admissions) divided by the stratum measure of size (i.e., 1980 total admissions for all hospitals in the stratum). The first weight to be assigned, therefore, was the inverse of this, or the stratum measure of size divided by the hospital measure of size. The probability of selecting a given record was given by the within sampling rate (WSR). For example, if the WSR was 1/150, every record had one chance in 150 of being selected and so represented 150 records in the hospital. The second weight to be applied, therefore, was the inverse of the WSR (in this instance 150).

As noted, the remaining weights were used to adjust for records eliminated in the process of locating and abstracting the records on the original sampling list. In making these adjustments two types of nonavailable records were identified. The first consisted of records which could not be located, were incomplete once found, or were removed subsequently by the Medical Records Department. This constituted the true "missing" category. The second category consisted of records which were "out-of-scope," i.e., which were for populations not covered by the study. These included pediatric cases (including newborns), normal deliveries, patients occupying subacute beds, and patients not discharged in 1981. These were treated differently in the weighting. For true missing records, the weight for nonmissing records was inflated by a proportionate factor to compensate for the missing record. For out of scope records the weight was not inflated since these records were not part of the population the study wished to estimate. Calculation of the three adjustment weights is described further below.

Using the WSR determined for the hospital, a Sampling Plan Record (see Appendix G) was generated for each hospital. This list was used for record identification and selection. Of the records on the list the hospital medical records department pulled all that could be located. The pulling of out of scope records was not required; however, medical records personnel were required to indicate which records were not pulled for these reasons. The third weight corrected for the proportion of records unavailable either because they were out of scope or because they were not located and was equal to the total number of records on the original sampling list, excluding those which were out of scope, divided by the number of records left after excluding both out of scope and missing records. By excluding out of scope records from the numerator and denominator, the weight ensures that the obtained records will be weighted only to the population of interest. In addition, the exclusion of missing records in the denominator has the effect of ensuring that the nonmissing records are weighted up to compensate for the missing records.

The next two weights applied to records eliminated in the abstracting process. Upon entering the hospital and receiving the records pulled by the medical records department, the abstractor was required to randomly order these on the Abstracting Record (presented in Appendix H). The abstractor then applied the Standardized Medreview Instrument to as many records as could be completed in the allotted time, eliminating any out-of-scope records that were found as the records on the list were abstracted. This produced two types of adjustment weights. The first was designed to adjust for the records which the abstractor could not complete in the allotted time and was equal to the total number of records to be abstracted divided by this number minus the number which the abstractor was unable to review. This weight, in effect, makes each record abstracted count also for the nonabstracted records.

The final weight applies only to the list of records that the abstractor was able to review and adjusts for the out of scope and missing records of this set. It is equal to the total number of records reviewed (excluding those that were found to be out of scope) divided by the number of records reviewed excluding both out of scope and missing records. As with the third weight adjustment, this weight adjusts the nonmissing records upward to compensate for the missing records, but first excludes out of scope records (since they are not part of the population which the study wished to estimate).

The final weight for a sampled record was the product of these five weights, namely:

$$\begin{aligned}\text{Record Weight} = {} & \frac{\text{strata MOS}}{\text{hospital MOS}} \\ & \times \frac{1}{\text{WSR}} \\ & \times \frac{\text{total sampled records - out of scope records}}{\text{total sampled records-(out of scope + missing records)}} \\ & \times \frac{\text{total records available for abstracting}}{\begin{array}{l}\text{available records - records not abstracted} \\ \text{because of lack of time}\end{array}} \\ & \times \frac{\text{total records reviewed - out of scope records}}{\text{total records reviewed -(out of scope + missing records)}}\end{aligned}$$

In addition to the weight applied to records, another weight was applied to the individual days of care. This weight allowed the days sampled from the stay to be weighted to represent the entire stay and was equal to:

$$\text{Day Weight} = \text{Record Weight} \times \frac{\text{total length of stay}}{\text{number of sampled days}}$$

b. Application of the Stage 1 and Stage 2 ratio estimators

The goal of the survey is essentially to estimate the following quantity (r), for the nation as a whole and for various subpopulations.

$$r = \frac{x}{y}$$

where:

x = the number of nonacute admissions or days
y = the total number of admissions or days

This quantity is termed a ratio estimate since both x and y are estimates derived for the sample. Since x is most likely correlated with y, in that there will be a greater number of nonacute admissions or days for any subpopulation having a greater number of total admissions or days, sampling precision is apt to be increased by increasing the precision with which y (total admissions or days) is estimated. This can be done if there is an independent source of data which gives the exact (or a more precise) value of y for the populations and subpopulations of interest.

In the present study more precise estimates of y, i.e., admissions or days of care, were available from the National Hospital Discharge Survey which sampled 226,585 patient records from over 400 hospitals in 1981. As noted previously, with pediatric cases (age <16 years) and normal deliveries excluded, the NHDS covers the same population as that represented in the present survey, except that the NHDS also included specialty hospitals (and provides no means of identifying and excluding such hospitals). To compensate for this, admissions and days of care for nonpediatric specialty hospitals were obtained from the AHA Annual Survey of Hospitals. These were then subtracted from the NHDS estimates for each of the subpopulations used in deriving the ratio estimators. The statistics from the AHA included a small portion of pediatric and subacute care which could not be reliably identified from the data, but this portion was not considered large enough to significantly affect the weights.

The NHDS and AHA data were used in the calculation of first and second stage ratio estimators for both admissions and days of care using a method adopted by the Current Population Survey and described in Hansen, Hurwitz, and Madow (1953). The first stage ratio estimator is designed to reduce the between hospital, i.e., between primary sampling unit (PSU), contribution to sampling variability by improving estimates of y based on selected hospital-level characteristics such as size or location. The second-stage ratio estimator is designed to reduce both the between-PSU and the within-PSU variances by improving estimates of y based on both hospital and patient characteristics.

The factors used in developing these ratio estimators were selected on the basis of their presumed relationship to nonacute utilization as evidenced in preliminary descriptive analyses, as well as their broad classificatory relevance for assuring the representativeness of sample results. Three factors, therefore, were used in developing the first stage ratio estimators: (1) bedsize (which was collapsed into four categories of 6 to 99 beds, 100 to 199 beds, 200 to 399 beds and 400 or more beds), (2) location, using the nine census divisions, and (3) ownership (classified as proprietary, nonprofit, and government nonfederal). Two factors were used in developing the second stage ratio estimators: (1) location, again using the nine census divisions and (2) whether the admission was medical or surgical, surgical being defined as any admission in which a UHDDS-defined Class 1 procedure was performed.

Ratio estimators were developed within these categories in the following manner. The final (second stage ratio) estimate is given by:

$$x'' = \sum_{a} \frac{x''_a}{y''_a} y_a$$

where Y_a is the independent estimate of total admissions or days in the ath division by medical/surgical status category (taken from the NHDS and adjusted with the AHA survey results); x''_a is the SMI estimate of nonacute admissions or days for the same subset; and y''_a is the SMI estimate of total admissions or days for that subset. Prior to calculating this quantity both x''_a and y''_a are adjusted using the first stage ratio estimation procedure as follows:

$$x''_a = \sum_{c} x_{ac} \frac{Z_c}{z'_c}$$

$$y''_a = \sum_{c} y_{ac} \frac{Z_c}{z'_c}$$

where x_{ac} is the SMI estimate of the number of nonacute admissions or days in the ath division by medical/surgical category and in the cth division by bedsize by ownership category, y_{ac} is the SMI estimate of total admissions or days for the same group, z'_c is the SMI estimate of total admission or days in the cth division by bedsize by ownership category and Z_c is the independent estimate (NHDS and AHA) of total admissions or days in the cth division by bedsize by ownership category.

The weights used for the final ratio estimate were derived as follows: Using the original weights, separate estimates of total admissions and days of care were made for each division by bedsize by ownership category from the SMI sample (z'_c) and the NHDS survey (Z_c), the latter being adjusted using AHA data. The fraction, $\frac{Z_c}{z'_c}$ was then multiplied by the original weight for admissions and days to obtain a first stage ratio estimation weight. Using this weight, total admissions and days of care were then estimated for each division by medical/surgical status category for the SMI (y'_a). The same quantities (Y_a) were then estimated from the NHDS sample (again adjusted using the AHA survey). A new weight $\frac{Y_a}{y'_a}$ was calculated and multiplied by the first stage ratio estimation weight. This produced the final (second stage ratio estimation) weight used in calculating the descriptive results in subsequent sections.

2. Determination of sampling variance

Many researchers have noted the difficulty of calculating the variance of sampling estimates from complex surveys. If direct computational procedures are available they may be too difficult and costly to calculate, even with the use of high-speed computers, for surveys with highly stratified, multi-stage cluster samples and elaborate estimation techniques. Furthermore, for some techniques used in sample survey designs, notably the use of controlled selection and sampling of only one first stage sampling unit from each stratum, the effect on the sampling variance cannot be calculated from the sample. Finally, if one wishes to go beyond descriptive analysis based on the estimation of population means, percentages, and totals to more complex analytical techniques (such as multiple regression), standard statistical techniques for assessing the significance of results do not apply directly.

Because each of these conditions prevailed in the present survey, analysis relied on the technique of within-sample replication of the estimate to derive a measure of its variance. The method used is taken from McCarthy (1979a, 1979b) and involves subsampling the "parent" sample to produce a number of pseudo-replicated estimates of any specific statistic. The precision of the statistic is then estimated from the variance of the replicated estimate by the following formula:

where $$s^2_{z'} = \frac{1}{n}\sum_{i=1}^{n}(z'_i - z')^2$$

n = the number of replicates

z' = the estimate of z based on the entire sample

z'_i = the estimate of z based on the ith subsample.

The variance of a number of types of statistics can be measured with this method including means, totals, regression coefficients and regression weights.

The replication technique used in the study is that which McCarthy terms "balanced half-sample replication." Using this method, one may obtain reasonably precise estimates of the sampling variance of the statistic with only 20 to 40 replicates. This is accomplished by efficient selection of replicates so that the between strata contribution to the estimates of variance is eliminated. These replicates "... have a variance algebraically identical with that of all possible replicates of the same character within the parent sample, and the same expected value as the variance of all possible replicates of the same character for all possible parent samples of the same design". (NCHS, 1966).

The replicates are constructed by first grouping all PSUs in the sample into homogeneous pairs. One member of this pair is then included in each replicate or "half-sample" in such a way that each member of the pair is included in half of the half-samples. To determine which member of the pair is selected in a given half-sample, one uses the method described by Plackett and Burman (1943-46) to construct a k x k orthogonal matrix of + and - signs, where k is a multiple of 4 and is greater than or equal to the number of strata, as in the following 8 x 8 matrix:

	pair 1	pair 2	pair 3	pair 4	pair 5	pair 6	pair 7	pair 8
Sample 1:	+	-	-	+	-	+	+	+
Sample 2:	+	+	-	-	+	-	+	+
Sample 3:	+	+	+	-	-	+	-	+
Sample 4:	-	+	+	+	-	-	+	-
Sample 5:	+	-	+	+	+	-	-	+
Sample 6:	-	+	-	+	+	+	-	-
Sample 7:	-	-	+	-	+	+	+	-
Sample 8:	-	-	-	-	-	-	-	-

If the sign in the matrix is +, the first member of the pair (determined by random assignment) is selected for the half sample. If the sign is negative, the second member of the pair is selected. The matrix above has the property that every pair of columns is orthogonal, which allows the cross product terms between pairs of PSUs to drop out over the entire set of half-samples.

The number of replicates (or half samples) which must be used depends upon the number of pairs that have been formed. As demonstrated by Gurney and Jewett (1975) there must be p^n replicates, where p is the number of units per stratum (in this case 2), and n is an integer value. One chooses the smallest value of n so that the following expression $(p^n-1)/(p-1)$ is equal to or greater than the number of pairs. To minimize the number of replicates needed in the present study, the balanced repeated replication technique was performed separately within each division, a tactic which also capitalized on the fact that the ratio estimation process was self-contained within each division.

Pairs were formed by grouping hospital strata similar in bedsize (and preferably with the same ownership status) within a division. If there was an odd number of hospitals within a division, the last pair was formed by splitting the admissions of the last sampled hospital into two random halves and letting each set of admissions comprise one half of the pair. The division with the greatest number of PSUs contained 41 hospitals which were formed into 21 pairs. As given by the formula above, 2^5 or 32 half samples were required to provide replicated estimates for this division, since n=5 is the smallest integer that can be used in the expression above to produce a value greater than 21:

$$(2^5-1)/(2-1) = 31$$

Although some divisions required as few as 8 replications to derive estimates of sampling variance, 32 replicates were used for each division for computational simplicity.

The balanced repeated replication technique was used to estimate sampling variance for the descriptive results. These descriptive results were essentially estimates of the proportion of nonacute admissions or days for the population as a whole and for various subpopulations. Since each proportion is a ratio estimate (i.e., of total nonacute days or admissions over total days or admissions), sampling variance is given by the following formula from Hansen, Hurwitz, and Madow (1953) where V_x^2 is the relvariance of nonacute admissions (or days), V_y^2 is the relvariance of total admissions (or days), and V^2_{xy} is the rel-covariance of these two quantities:

$$V^2_{(x/y)} = V^2_x + V^2_y - 2\, V_{xy}$$

The covariance term was estimated from the 32 half samples as follows:

$$S_{xy} = \sum_{j}^{32} \frac{(x_j - \bar{x})(y_j - \bar{y})}{32}$$

where x_j is the estimated number of nonacute admissions (or days) for the subpopulation in the jth half sample, y_j is the estimated number of all admissions (or days) in the jth half sample, x is the estimated number of nonacute admissions (or days) in the division as a whole, and y is the estimated number of all admissions (or days) in the division as a whole. The calculation of variance is a special case of the calculation of the covariance, with x_j and x substituted for y_j and y to obtain a variance estimate (S_{xx} or S^2_x) of x. The converse is done to obtain a variance estimate (S_{yy} or S^2_y) of y. Relvariances were then obtained by dividing the variance by the square of the estimate (i.e., $V^2_x = S^2_x/x^2$) and the relcovariance was obtained by dividing the covariance by the product of x any y (i.e., $V_{xy} = S_{xy}/xy$).

To estimate the sampling variance of a given ratio $\frac{x}{y}$, the balanced repeated replication technique was used in the manner described above to obtain the relvariances of x and y and the rel-covariance for each division. Once these values were calculated for each division, they were then summed across divisions to obtain national sample values of the relvariance of x, the relvariance of y, and their relcovariance. To obtain an exact estimate of sampling variance these terms were then applied in formula (1) above to estimate the variance of the ratio in the national sample. A method for generalizing these estimates of sampling variance to other quantities calculated in the course of the analysis is described below.

3. Generalization of sampling variances

The balanced repeated replication method of estimating sampling variance is costly and prohibitive when a large number of statistics are to be reported. As a result, the study used an algorithm developed by Steinberg (1965) for generalizing the relvariances obtained using the balanced repeated replication method. To use this algorithm, the balanced repeated replication method was first used to calculate relvariances ($V^2_{x'}$) for twenty estimated totals (x'), such as the total number of nonacute admissions in the Southern region or the total number of admissions for Medicare patients. These twenty x' and the associated relvariances were then input into the algorithm (presented in Appendix J) which finds the a and b (see below) which minimize the squared relative residuals of $V^2_{x'}$:

$$V^2_{x'} = a + \frac{b}{x'}$$

Once a and b are determined, they provide a formula for determining the relvariance, $V^2_{x'}$, of any new x' that may be estimated as part of the analysis.

Separate generalization functions were estimated for admissions and days. The following function was obtained for admissions:

$$V^2_{x'} = .0001956 + \frac{11{,}643}{x'} \qquad (3)$$

while the function for days was estimated as:

$$V^2_{x'} = .0007454 + \frac{192149}{x'} \quad (4)$$

These functions are used in the following manner to determine the sampling variance of a given proportion: Suppose the proportion of nonacute admissions for patients 65 years of age or over is estimated as 406,000 nonacute admissions in a total of 10,150,000 admissions for this age group ($\underline{x}$ = .04). Formula (3) above, the generalization function for admissions, would be used to calculate, first, the relvariance of the estimated nonacute admissions and, then, the relvariance of the estimated total admissions as follows:

$$V^2_{x'} = .0001956 + \frac{11643}{406{,}000} = .03063$$

$$V^2_{y'} = .0001956 + \frac{11643}{10{,}150{,}000} = .00134$$

These provide the terms V^2_x and V^2_y respectively in the before-mentioned formula for calculating the variance of a ratio estimator:

$$V^2_{(x/y)} = V^2_x + V^2_y - 2\ V_{xy}$$

No covariance term, however is available from the generalization function but, a reasonable approximation for $V_{(x/y)}$ is given by subtracting the relvariance of y' from the relvariance of x' as follows:

$$V^2_{(x/y)} = V^2_x - V^2_y$$

This value is then multiplied by the square of the ratio itself to obtain the variance (rather than relvariance) of the estimate. The square root of the result provides the standard error of the estimate. In the example given above, the relvariance of the ratio (as estimated by the generalization function) is:

$$V^2_{(x/y)} = .03063 - .00134 = .02929$$

The variance of the estimate is:

$$S^2_{(x/y)} = \frac{406,000}{10,150,000}^{2} \text{ X } .02929 = .00004686$$

And the standard error of the estimate is:

$$S_{(x/y)} = \sqrt{.00004686} = .00685$$

Therefore, in the hypothetical example above, the standard error for the ratio, .04, is ±.0069 and the 95% confidence limit is ±.0138.

This function was used in estimating most of the standard errors given in the following sections. However, formula (1), using the values obtained in the balanced repeated replication process, was used for certain values which are noted in the tables.

C. RELIABILITY AND VALIDITY OF THE NATIONAL SURVEY DATA

The development and field testing of the methodology, described in Chapter III, incorporated a comprehensive assessment of reliability as well as various checks on validity. Two further assessments were made for the national survey data: the first was a test of inter-rater reliability, the second was a study of the logic and internal consistency of abstractors' evaluations as to the acuteness or nonacuteness of admissions and days of care. These will be described in turn.

1. Assessment of Inter-Rater Reliability

Since the Standardized Medreview Instrument was extensively revised for the national survey, an additional assessment of reliability was made which allowed us to measure its performance under actual field conditions. This assessment entailed reabstracting all records in seven of the survey hospitals. These were randomly selected from each of seven mutually exclusive and exhaustive sets of hospitals, from seven different areas of the country.

For each of the selected hospitals an abstractor was selected to return to the hospital and re-evaluate the same set of patient records that had been reviewed previously. This abstractor was one who had participated in the study, but was not the one who had originally reviewed the records of the selected hospital.

In all, 124 admissions and 1245 days of care were reabstracted. Agreement rates are presented below:

		DAYS			ADMISSIONS		
		Reviewer 1:			Reviewer 1:		
		NONACUTE	ACUTE		NONACUTE	ACUTE	
Reviewer 2	NONACUTE	33	12	45	5	2	7
	ACUTE	4	1,196	1,200	1	116	117
		37	1,208	1,245	6	118	124
	Percentage Agreement[1]	= 98.7%			= 97.6%		
	Kappa[2]	= 79.8%			= 75.7%		

---

[1] Percentage Agreement equals the percent of cases both abstractors call acute plus the percent of cases both reviewers call nonacute. Referred to as "hits" by some.

[2] Kappa (Cohen, 1976) corrects Percentage Agreement for the percentage of agreement that would be expected by chance alone given the marginal distributions, and is considered the most definitive measure of agreement.

As can be seen, percentage agreement is high (99% for admissions and 98% for days), but this is due to the low rates of nonacute admissions and days (around 5%) found in the data set. Low occurrence rates ensure a high overlap in reviewer judgments due to chance alone. In general, it is difficult to achieve high levels of real (as opposed to chance) agreement with phenomena having low rates of occurrence. The small range of variation in the levels of acute admissions and days attenuates possible correlations with other variables (including correlations of one rater's judgments with those of another rater). Nevertheless, interobserver agreement in the reabstracting study was respectable and compared favorably to the measured reliability of other instruments for measuring inappropriate agreement. The value of Kappa found in the study (which corrects the percentage agreement observed in the sample for the percentage that would be expected by chance alone) was .76 for admissions and .80 for days. Both these values were significantly greater than zero ($p<.001$).

2. Assessment of the Validity of Reviewer Evaluations

At the conclusion of the data collection effort, reviewer evaluations of admissions and days of care were tabulated. Exhibit IV-7 presents an unweighted breakdown of the admission evaluations. Of the admissions, it can be seen that 63% were judged acute because they met a specific admission criterion. Of the remaining admissions, 33% were justified on other grounds: 23.9% did not meet a criterion, but were for a Class 1 procedure that required an inpatient setting, 5.1% met an extension of a criterion, and 4.0% met no criterion, but were judged acute on the basis of record review. In all, 3.7% of the admissions were judged nonacute: 181 (or 3.2%) met no criterion and 30 (or 0.5%) met a criterion which was overridden by the reviewer. If extensions of criteria are counted as a form of override, it can be seen that overrides in the direction of justifying an admission were more frequent (at 9.1%) than overrides in the opposite direction (at 0.5%). In all, the percentage of evaluations which were reviewer overrides was almost 10%.

EXHIBIT IV-7

RESULTS OF THE ADMISSION EVALUATION

(unweighted)

	n	%
Admission Justified:		
Criterion Met	3,628	63.3%
For Class 1 Procedure	1,371	23.9%
Criterion Met By Extension	292	5.1%
No Criterion Met, but Record Review Supports	230	4.0%
Admission Not Justified:		
No Criterion Met	181	3.2%
Criterion Met, but Record Review Does Not Support	30	0.5%
TOTAL	5,732	100.0%

n = number of cases in the sample

% = percentage of cases in the sample

Exhibit IV-8 presents the level of service evaluations and level of service criteria for those stays where the admission was judged to be appropriate. Of the 22,913 days, level of service was judged to be appropriate 93.7% of the time: In 89.2% of the cases a criterion was met directly, another 2.3% met an extension of a criterion and 2.2% were justified on the basis of record review. These figures decline slightly when considering only days where the need for continued hospitalization was also justified. For 6.3% of the days the level of services was not appropriate. Of these 5.6% met no criterion. For another 0.7% of the days the criterion was overridden by the reviewer. As with the admission evaluation, overrides in the direction of justifying the level of service exceeded those in the opposite direction, by 4.5% to 0.7%.

Exhibit IV-9 presents similar information for the need for continued hospitalization review. On 93% of the days, the patient was judged to require continued hospitalization. Of these, 87% met a criterion directly, 2% met an extension of a criterion and 4% were justified on the basis of record review. These figures each decline slightly when considering only days where the level of service is also appropriate. On a total of 6.8% of the days the patient was not considered to need continued hospitalization. For 6.2% of these, the day met no criterion and on 0.6% the day met a criterion and was overridden. Again overrides in the direction of justifying the need for continued hospitalization exceeded those in the opposite direction by 6.1% to 0.6%.

The instrument development phase of the study, described in Chapter III, was performed with the goal of selecting the most complete and valid criteria set possible for assessing the acuteness of care provided at the hospital level. Accordingly, a major component of the data analysis prior to the actual examination of substantive results was to examine how the criteria were used by abstractors to determine whether the rates of nonacute use reported did, in fact, adhere to the standards established at the outset of the study. This was particularly the case since the reviewer overrides of the criteria were so greatly in favor of justifying the admission or day.

EXHIBIT IV-8

LEVEL OF SERVICE EVALUATIONS AND LEVEL OF SERVICE CRITERIA FOR JUSTIFIED ADMISSIONS

(unweighted)

	FREQUENCY	%	FREQUENCY	%
LEVEL OF SERVICE JUSTIFIED:	(includes all days in justified admissions)		(includes only days where both continued stay and level of service are justified)	
Criterion Met	20,429	89.2%	20,137	95.4%
Criterion Met by Extension	524	2.3%	487	2.3%
No Criterion Met, but Record Review Supports				
LEVEL OF SERVICE NOT JUSTIFIED:				
No Criterion Met	1,281	5.6%	0	0
Criterion Met but Record Review Does Not Support	164	0.7%	0	0
TOTAL DAYS FOR JUSTIFIED ADMISSIONS	22,913	100.0%	21,109	100.0%

EXHIBIT IV-9

CONTINUED STAY EVALUATIONS AND CONTINUED STAY CRITERIA FOR APPROPRIATE ADMISSIONS

(unweighted)

	FREQUENCY	%	FREQUENCY	%
CONTINUED STAY JUSTIFIED:	(includes all days in justified admissions)		(includes only days where both continued stay and level of service are justified)	
Criterion Met	19,958	87.1%	19,759	93.6%
Criterion Met by Extension	474	2.1%	458	2.2%
No Criterion Met but Record Review Supports	922	4.0%	892	4.2%
CONTINUED STAY NOT JUSTIFIED:				
No Criterion Met	1,426	6.2%	0	0
Criterion Met, but Record Review Does Not Support	133	0.6%	0	0
TOTAL DAYS FOR JUSTIFIED ADMISSIONS	22,913	100.0%	21,109	100.0%

An intensive examination was performed of reviewer's decisions concerning the appropriateness of admissions. Because the results of this analysis were largely favorable, only a brief examination of daily judgments was made. These investigations are summarized in turn below. A detailed discussion of the methods and findings of this evaluation and the decisions made can be found in Appendix K.

a. Examination of Reviewers' Admission Judgments

Because an intensive examination of every admission in the dataset was prohibitive, the validity analysis proceeded in a focused manner, successively targeting those classes of admissions about which there were most likely to be questions. The first step of the analysis was to examine admissions that met no criterion but were justified because they were for operations or procedures which require an acute care setting. This was followed by evaluations of admissions justified by an extension of a criterion and those justified on the basis of record review. Thirdly, all admissions judged nonacute were examined. Finally a focused analysis of 1,425 admissions which met a criterion and were judged acute was undertaken. These analyses are described in turn below.

i. Examination of Admissions for Procedure

As noted, the methodology for the study made no attempt to evaluate whether or not admissions for elective procedures were warranted given the patient's condition. Rather, the evaluation was restricted to a judgment as to whether the procedure itself required the acute care setting. Therefore, our examination of admissions justified for the performance of a procedure was concerned only with whether the procedures performed were those best performed in an inpatient setting.

To perform this investigation, we selected all admissions which fell into this category and individually examined the principal procedure and any secondary procedures which were performed. A distribution of principal procedures (coded using ICD-9-CM nomenclature) is shown in Appendix K, section 2a. These procedures were compared with lists of procedures appropriate for an inpatient setting taken from several sources. Any procedures about which there was doubt were referred to the physician consultant.

At the end of this examination we reversed only 9 of the 1,371 admissions falling in this category. Some of these were for procedures (such as tooth extractions) which should be performed in an outpatient setting. Most, however, were admissions in which the patient was scheduled for an elective procedure but the procedure was cancelled because of patient noncompliance or controllable delays in service.

There were also 93 admissions (representing 6.8% of the total number of admissions for a procedure) which -- although not performable in all outpatient settings -- could be performed in an ambulatory surgical unit, given the availability of such a unit. These admissions were for D & Cs, tubal ligations and so forth. Since medical practice is rapidly evolving in this area and there is widespread variation among hospitals and geographic locations in the availability of ambulatory surgery units, these admissions were not classified as unjustified. They were, however, specially categorized so that they could be identified in the analysis.

ii. Examination of Admissions Justified as Extensions of Criteria on the Basis of Record Review

In applying the Standardized Medreview Instrument, if an admission did not meet a criterion, the abstractors then reviewed all documentation pertaining to the need for hospitalization to ascertain whether or not the patient actually did require acute level services. If so, the reviewer was permitted to exercise an override judgement. If the case would have met a criterion had the criterion been broader in scope, the reviewer indicated that the need for admission was supported by a logical extension of one of the criteria. If the admission was justified on other grounds (e.g., age or comorbidity) the reviewer indicated that the admission was justified on the basis of record review. In both cases, the reviewer was required to enter a written explanation on the abstract.

A detailed examination of these evaluations was made, which included a review of the original abstract and written explanation and, in some cases, submission of the evaluation to the physician consultant. Exhibit IV-10 presents the results of this analysis. In all, 52 (10%) of the 523 reviewers' evaluations were reversed, and 471 (90%) were supported. In the majority of instances (375 in all) the admission was justified on the basis of age or

EXHIBIT IV-10

ANALYSIS OF ADMISSIONS JUSTIFIED BY EXTENSION OR RECORD REVIEW

(unweighted)

	TOTAL
REVIEWER JUDGMENT SUPPORTED:	
Principal or Secondary DX Met Criterion	8
Admission Justified for Procedure	58
Admission Justified on the Basis of Age, Comorbidity, or Definitive Treatment	375
Admission Supported by Physician Review	30
REVIEWER JUDGMENT REVERSED:	
Reversed on the basis of Abstract Review	4
Reversed by Physician Review	48
TOTAL	523

comorbidity or because the patient was receiving definitive treatment of a type available only in a hospital. In some cases the principle diagnosis or a secondary diagnosis actually met a criterion, and in 58 cases the patient was appropriately admitted for performance of a procedure. Seventy-eight cases were more complex and were submitted to the physician consultant. Of these, the physician supported the reviewer in 30 instances and reversed the reviewer's decision in 48.

iii. Examination of Admissions Evaluated as Nonacute

This part of the investigation was concerned with admissions evaluated by the reviewer as not justified, either because no criterion was met or because the record review failed to support the need for admission, even though a criterion was met. The same type of case by case examination was made for this class of admissions as described above. The results of this analysis are shown in Exhibit IV-11. In 22 of these admissions (10% of the total) the reviewer's evaluation was not supported, that is, we considered the admission to be justified. In 189 of the cases, we concurred with the reviewer that the admission was not justified.

iv. Examination of Admissions Meeting a Criterion

The previously described analyses focused successively on (1) admissions which were justified on the basis that a Class 1 procedure was required, (2) those justified as an extension of a criterion or on the basis of record review, and (3) those judged nonacute. The largest class of admissions in the sample was, of course, those which, in fact, met a criterion. Although it was not feasible to examine every one of these admissions, we adopted a systematic strategy to identify possible misinterpretations of the Standardized Medreview Instrument criteria set.

The first step in this process was to examine all admissions having the same primary diagnosis as one of the admissions judged to be nonacute at the conclusion of the above investigations. A total of 850 admissions met this condition. These are presented by primary diagnosis (given in five digit ICD-9-CM codes) in Exhibit IV-12. As can be seen in Exhibit IV-12, the most frequently occurring diagnoses were 162.90 (Lung Cancer) with 3% of the cases;

EXHIBIT IV-11

ANALYSIS OF ADMISSIONS WHERE NO CRITERION MET OR RECORD REVIEW FAILS TO SUPPORT

	NO CRITERION MET	RECORD REVIEW DOES NOT SUPPORT
REVIEWER JUDGMENT NOT SUPPORTED:		
Admission Meets Criterion	4	2
For Class 1 Procedure	1	1
For Procedure That Could be Outpatient If One Available	7	0
Rural Area: No Outpatient Facility Available	1	0
Overturned by Abstract Review	3	0
Overturned by Physician Review	3	0
REVIEWER JUDGMENT SUPPORTED:		
Supported by Abstract Review	159	27
Supported by Physician Review	3	0
TOTALS	181	30

EXHIBIT IV-12

DISTRIBUTION BY IDC-9-CM DIAGNOSIS CODE OF ACUTE ADMISSIONS WITH THE SAME PRIMARY DIAGNOSIS AS A NONACUTE ADMISSION

PRINCIPAL DIAGNOSIS (ICD-9-CM)	FREQUENCY	PERCENT
034.00	3	0.353
079.90	4	0.471
153.90	7	0.824
154.00	2	0.235
154.10	4	0.471
162.90	28	3.294
174.90	11	1.294
185.00	10	1.176
189.00	1	0.118
191.30	1	0.118
202.80	1	0.118
211.30	1	0.118
239.00	1	0.118
250.00	20	2.353
250.60	3	0.353
296.20	4	0.471
300.00	2	0.235
300.40	13	1.529
303.00	14	1.647
308.90	1	0.118
310.90	7	0.824
311.00	12	1.412
346.90	4	0.471
354.00	1	0.118
401.90	20	2.353
413.90	55	6.471
414.00	67	7.882
437.90	4	0.471
455.80	1	0.118
466.00	20	2.353
470.00	1	0.118
473.00	1	1.118
487.10	4	0.471
490.00	5	0.588
493.90	19	2.235
496.00	50	5.882
530.10	1	0.118
530.30	1	0.118
532.90	16	1.882
533.30	1	0.118
533.90	4	0.471
535.00	5	0.588
535.50	12	1.412
535.60	2	0.235
536.80	2	0.235
536.90	3	0.353
553.30	9	1.059
556.00	2	0.235
558.90	34	4.000
560.39	1	0.118
562.10	8	0.941

EXHIBIT IV-12 (continued)

PRINCIPAL DIAGNOSIS (ICD-9-CM)	FREQUENCY	PERCENT
562.11	15	1.765
564.10	6	0.706
564.90	3	0.353
565.00	2	0.235
569.30	2	0.235
571.20	9	1.059
571.49	1	0.118
574.10	15	1.765
575.10	3	0.353
577.00	18	2.118
590.10	4	0.471
595.90	7	0.824
599.00	34	4.000
599.70	7	0.824
600.00	27	3.176
604.90	3	0.353
611.72	1	0.118
614.20	3	0.353
614.90	15	1.765
618.00	3	0.353
620.10	1	0.118
620.20	8	0.941
621.60	1	0.118
623.80	1	0.118
626.80	8	0.941
682.40	1	0.118
682.70	5	0.588
696.10	1	0.118
717.90	1	0.118
722.40	1	0.118
724.20	8	0.941
724.40	8	0.941
729.10	4	0.471
733.00	4	0.471
733.10	4	0.471
733.82	1	0.118
780.20	17	2.000
780.40	1	0.118
780.60	4	0.471
783.00	1	0.118
784.00	2	0.235
787.00	2	0.235
788.20	2	0.235
789.00	37	4.353
789.30	2	0.235
805.20	3	0.353
807.01	2	0.235
846.00	9	1.059
846.90	4	0.471
847.00	6	0.706
847.20	4	0.471
847.90	3	0.353
920.00	7	0.824
959.10	2	0.235
998.10	4	0.471

250.00 (Adult Onset Diabetes Mellitus without complications) with 2%, 401.90 (Essential Hypertension) with 2%, 413.90 (Angina Pectoris) with 6%, 414.00 (Chronic Ischemic Heart Disease) with 8%, 466.00 (Acute Bronchitis) with 2%, 496.00 (COPD) with 6%, 558.90 (Noninfectious Gastroenteritis and Colitis) with 4%, 599.00 (Urinary Tract Infection) with 4%, and 789.00 (Abdominal Pain) with 4%.

These diagnoses were not concentrated among a few criteria but covered close to the entire range of criteria on the instrument. Therefore, it did not appear that these admissions were being justified by a small number of possibly weak criteria. A more extensive examination of these admissions was undertaken to ascertain whether there were, in fact, conditions which distinguished them from the nonacute admissions identified in the previous investigation. A detailed summary of each admission that was reviewed is given in Appendix K. The following information highlights the major findings of this investigation.

A first glance at the 850 cases targeted for investigation showed that 251 of them (or 30%) met a criterion for laboratory abnormality (criteria 1 to 28 on the instrument). These laboratory values are quite stringent, and are generally considered sufficient to warrant admission without question in their own right. A significant number of what might be considered weaker diagnoses met these laboratory criteria, including 80% of the patients with Adult Onset Diabetes and no coded complications (code 250.00), 65% of the patients with Acute Bronchitis (code 466.00), 52% of the patients with COPD (code 496.00), 56% of the patients with Acute Pancreatitis (code 577.00), and 85% of the patients with Urinary Tract Infection (code 599.00).

Further study was then made of the 599 admissions having one of the targeted primary diagnoses but not meeting a laboratory abnormality criterion. Each admission was examined individually in a review that took into account the secondary diagnoses, procedures performed, the level of services throughout the stay, patient age, length of stay, and the reviewer's written notes. By the conclusion of this investigation, we had identified 25 admissions for which the evaluation was reversed. This amounted to 2.9% of the 850 cases that had been targeted for investigation. Thus, we found that, overwhelmingly, although the primary diagnoses for these admissions were the same as for those admissions previously judged nonacute, there were a number of factors which supported the need for these admissions.

To follow up on this investigation, we selected another set of admissions for which the primary diagnosis was *not* the same as one of the nonacute admissions. We decided to investigate all other admissions justified on the basis of the criteria (1) "Pain, severe and incapacitating," of which there were 245 cases; (2) "Functional impairment," of which there were 71 cases; and (3) "Sensory motor deficit" of which there were 19 cases. (The latter category was selected as being a representative criterion from among the criteria concerned with abnormal signs.) In addition, we examined (1) all other admissions for Back Pain, Hypertension, and Diabetes Mellitus that had not been examined before (codes 840.90, 843.80, 844.10, 845.09, 848.80, 401.10, 250.00-250.99) of which there were 9 cases; (2) all admissions for Nonpsychotic Mental Disorders (codes 300.00 to 316.99) which did not meet the psychopathology criterion of which there were 69 cases and (3) all admissions which met the criterion for drug or alcohol abuse, but which were not placed in an alcohol dependency unit, of which there were 32 cases. In addition, two special classes of patients were examined. The first was 181 patients over 74 years old who were discharged to subacute care facilities or home care. The second class of special admissions we examined were 46 terminal cancer patients, i.e., all patients with malignant neoplasms (ICD-9-CM codes 140-208) who died in the hospital. This was done to assess whether reviewers were justifying admissions for terminal patients on humanitarian grounds that did not meet any of the criteria. The results of this investigation were similar to the previous one. In isolated instances, admission decisions were reversed. Generally, however, admissions met the criterion based on both diagnoses and the provision of acute care services and definitive treatment (at least in the initial few days of the stay).

In all, 575 additional admissions were examined in the second phase of the investigation. These, in addition to the 850 examined previously, amounted to a total of 1425 admissions. Of these, 34 admissions (2.4%) previously judged appropriate by the reviewer were reversed. In 14 of these we lacked information on level of services at the beginning of the stay since these days were not among the ones abstracted. This information could easily have been crucial in evaluating the appropriateness of the admission. To be conservative, however, we overrode the reviewer's judgement in order to get an upward bound figure on the possible degree of reviewer bias.

v. Summary

Exhibit IV-13 presents the result of our investigation of reviewers' evaluations. Of 5478 admissions that had originally been judged acute, 110 of the evaluations were reversed as a result of the investigation (34 were admissions that had previously met a criterion, 9 were admissions for a Class 1 procedure, 31 were admissions that had met an extension of a criterion, and 36 were admissions that had been judged acute on the basis of record review). It can be seen that a greater proportion of the admissions that were justified by extension or by record review were overturned than those which had met a criterion. Thirty-one out of 292 (11%) of those originally justified by extension and 36 out of 230 (16%) of those originally justified by record review were reversed, compared with 34 out of 1425 (2.4%) of those that had met a criterion. This, no doubt, reflects the more controversial status of admissions which do not definitively meet a criterion.

Of the 211 admissions originally judged nonacute, 21 or 10% of the evaluations were reversed. As a result, by the end of the investigation, 299 or 5.2% of all admissions (unweighted) were determined to be nonacute, as opposed to the 211 or 3.7% originally evaluated as nonacute. Thus, in our investigation we did find a bias toward understating the rate of nonacute hospital admissions. This bias was slight, however; in general, the reviewers' evaluations resisted a fairly aggressive effort to identify misinterpretation of the criteria.

In the following analyses we used the 299 nonacute admissions identified above to estimate the rates of nonacute admissions in the population. It should be noted that only 1425 of the 3627 admissions originally justified on the basis of a criterion were reviewed, 2.4% of which were re-evaluated as nonacute. If the same rate of reversals were found for the remaining 2203 admissions, this would make another 52 admissions nonacute, or a total of 6.1% overall. We thought it unlikely, however, that such a high rate of reversals would occur for these admissions. The 1425 admissions that were reviewed were those targeted as the most potentially subject to misuse (e.g., admissions with the same primary diagnosis as another admission judged to be nonacute, admissions justified on the basis of the "Pain" criterion, and so forth). These evaluations would, therefore, be more apt to be reversed than those for the remaining admissions.

EXHIBIT IV-13

SUMMARY RESULTS OF THE REVIEW OF ABSTRACTOR EVALUATIONS

ORIGINAL EVALUATION	FINAL EVALUATION		
Admission is Acute:	Acute	Nonacute	Total
Criterion Met			
Admission Examined	1391	34	1425
Admission Not Examined	2203	-	2203
For Procedure	1362[1]	9	1371
Extension of Criterion	261	31	292
Record Review Supports	194	36	230
	5410	110	5521
Admission is Nonacute:			
No Criterion Met	19	162	181
Criterion Over-ruled	3	27	30
	22	189	211

---

[1] Includes 93 admissions for procedures that could be in an outpatient surgical unit if one is available.

Additionally, in making our re-evaluations of the appropriateness of the admission, we no longer had all the information originally available to the reviewer. Therefore, we undoubtedly reversed some judgments that would not have been reversed had the medical record been available to us. We felt that the undetected nonacute admissions among the remaining admissions approximately balanced the evaluations we may have wrongly reversed.

This number is lower than that found in other studies using other methodologies (e.g., Gertman and Restuccia, 1981). The judgment of what constitutes inappropriate hospital use is, of course, subject to differences of opinion depending upon the medical and social philosophy of the individual. Even if a definition can be agreed upon, it is difficult to operationalize the concept in terms of objective and specific criteria. As a result of our investigations, however, we feel that the reviewers' evaluations adhere to the principles established by the criteria set and that the criteria set itself reflects established norms of medical care.

b. Investigation of Reviewers' Evaluations of Level of Service and Need for Continued Hospitalization

The investigation of reviewer evaluations of individual days was less intensive than that for admissions. Whereas the data set contained much supplementary information (e.g., diagnoses, procedures performed, level of service) useful in corroborating the reviewers' admission evaluation, this type of semi-redundant information was not available for the individual days of care. In addition, the more extensive evaluation of admissions had confirmed that the reviewers were for the great part exercising their judgment in a manner consistent with the logic of the instrument.

To evaluate the consistency of reviewers' judgments concerning level of service and need for continued hospitalization we, therefore, confined our analysis to the following. The first step of our evaluation was to examine specifically the level of service and need for continued hospitalization criteria that were chosen by extension. To do this a random sample of 100 days was selected for examination from among the 364 in which the reviewer had chosen both the level of service and continued hospitalization criterion by override (i.e., either by extension or on the basis of record review). The original abstract was pulled and the notes examined. Our review supported the abstractor's evaluation in all but one of these cases.

Because the criterion "General Observation" was used so frequently as a level of service criterion (14% of days when a level of service criterion was met and 23% of the days when one was chosen by extension), we performed a special examination of this criterion. Exhibit IV-14 presents the distribution of patient stays by the number of times in the stay that observation was checked as a criterion for each type of admission. As can be seen, general observation was rarely used more than the recommended two days for any type of admission. For admissions which met a criterion, only 4% of the stays used general observation to justify the level of services on more than two of the abstracted days, and only 2% had general observation on more than three days. For the remaining admission types, this figure was even lower.

EXHIBIT IV-14

DISTRIBUTION OF PATIENT STAYS BY NUMBER OF DAYS OF GENERAL OBSERVATION

Total Days of General Observation Per Stay	Admission Justified by Criterion		Admission Justified for Procedure		Admission Justified by Extension		Admission Justified by Record Review	
	n	%	n	%	n	%	n	%
0	2558	70.5	861	62.6	196	66.9	186	80.5
1	465	12.8	1111	18.2	43	14.7	21	9.1
2	464	12.8	226	16.4	45	15.4	20	8.7
3	80	2.2	23	1.7	6	2.0	4	1.7
4	33	0.9	9	0.7	1	0.3	0	0.0
5	22	0.6	4	0.3	0	0.0	0	0.0
6	5	0.1	1	0.1	2	0.7	0	0.0
7	2	0.1	0	0.0	0	0.0	0	0.0
TOTAL	3629	100.0	1374	100.0	293	100.0	231	100.0

V. STUDY RESULTS

V. STUDY RESULTS

After analyzing sample representativeness and the reliability and validity of reviewer evaluations as described in Chapter IV, the study was in a position to estimate rates of nonacute hospital utilization and to investigate the factors associated with such utilization. The results of the analysis are presented in this chapter.

A. NATIONAL ESTIMATES OF NONACUTE UTILIZATION

The national rates of nonacute utilization estimated in this study are presented in Exhibit V-1 for admissions and Exhibit V-2 for days. As shown in Exhibit V-1, the survey found a total of 1,925,000 or 5.7% of the admissions to be nonacute, i.e., admission to an acute care facility was not warranted. The 95% confidence interval for this figure is 1.2%. Of those admissions judged to be appropriate, 20.8% were specifically for elective procedures, and the remaining 73.5% were admissions that met a criterion or an extension of a criterion or were justified by record review. Of the 7,088,000 admissions for elective procedures, 577,000 or 8.1% could have been performed in an ambulatory surgery unit had one been available. This amounts to 1.7% of all admissions.

Exhibit V-2 shows that overall 12.0% of the hospital days were nonacute (with a 95% confidence interval of 2.0%). Of these, 4.2% were nonacute because the admission itself was nonacute (since all days in a nonacute admission were automatically considered nonacute). Another 7.8% of the nonacute days occurred in stays where the admission itself was justified. When only necessary admissions are considered, 8.1% of the 240,482,000 days occurring in these stays were nonacute.

Some statistics in the literature concerning the magnitude of unnecessary hospital use in U.S. hospitals are higher. Many are based on relatively unstructured methods for measuring inappropriate utilization (e.g., Berg et al, 1969) and all are based on small and nonrepresentative samples. The only published statistics based on a structured method for measuring nonacute

EXHIBIT V-1

NATIONAL ESTIMATES OF THE RATE OF NONACUTE HOSPITAL ADMISSIONS

	N (in 1,000s)	r	s.e.
ACUTE ADMISSIONS			
For Procedure	7,088	20.8%	0.75%
(Possible Ambulatory Surgery Unit)	(577)	(1.7%)	0.24%
(Definite Inpatient Procedure)	(6,511)	(19.1%)	0.73%
Other	24,988	73.5%	0.82%
NONACUTE ADMISSIONS	1,925	5.7%	0.60%e
	34,001	100.0%	

N = estimated admissions

r = ratio estimate of % admissions in each category over total estimated admissions

s.e. = standard error of r (estimated using the generalization function unless otherwise denoted by e, in which case standard error was calculated directly using the variances and covariances generated by the balanced repeated replication process).

EXHIBIT V-2

NATIONAL ESTIMATES OF THE RATE OF NONACUTE HOSPITAL DAYS

	N (in 1,000s)	r	s.e.
ACUTE ADMISSIONS			
Acute Days	220,883	88.0%	1.08%
Nonacute Days	19,599	7.8%	0.89% e
NONACUTE ADMISSIONS			
Nonacute Days	10,467	4.2%	0.66% e
TOTAL DAYS	250,949	100.0%	

N = estimated days

r = ratio estimate of % days in each category over total estimated days

s.e. = standard error of r (estimated using the generalization function unless otherwise denoted by e, in which case standard error was calculated directly using the variances and covariances generated by the balanced repeated replication process).

utilization similar to the Standardized Medreview Instrument are reported by Gertman and Restuccia (1981) who applied the AEP or Appropriateness Evaluation Protocol (prior to the addition of objective admission criteria) at Boston University Medical Center's University Hospital. These researchers found that close to 30% of the days were inappropriate, a figure they noted was similar to rates they had found in a representative sample of 8000 patients from 41 community and teaching hospitals in Eastern Massachusetts. In the same study reviewers, using subjective judgment only, found about 10% of the admissions to be inappropriate.

Given the difference in both methodology and the population studied between the two studies, it is difficult to draw conclusions. In the present study, the percentage of unnecessary admissions was found to be as high as 28% in some hospitals and the percentage of unnecessary days was as high as 43%. For studies based on single institutions, then, differences in the magnitude of nonacute utilization found in this study versus those in Gertman and Restuccia's study are not unexpected. These differences might also be attributable to regional variation, since this study found the highest percentage of nonacute days in the Northeastern region where the Massachusetts hospitals studied by Gertman and Restuccia are located. Nevertheless, in this study, even in the Northeast the percentage of nonacute days (counting days in both appropriate and inappropriate admissions) was only half of that found by these investigators.

Although the AEP and the SMI are quite similar in their logic and application, there are methodological differences which could contribute to the observed difference in rates of nonacute utilization. The AEP requires that a day meet only a patient condition criterion or a level of service criterion in order to be considered appropriate. With the SMI the patient must meet both a condition criterion and a level of service criterion. This should make the SMI more stringent, however, and increase the rates of nonacute utilization found with the SMI compared to those found with the AEP. The other difference between the two instruments is the number of criteria. The SMI has 56 day-level criteria compared with 27 for the AEP, and it has 117 admission level criteria compared with 18 for the most recent version of the AEP. A number of admissions and days of care which meet an explicit criterion on the SMI, therefore, may not on the AEP.

Although reviewers using the AEP may exercise the override option whenever an appropriate admission does not meet a criterion, there may nevertheless be a tendency with the AEP to deny the appropriateness of an admission or day that would be explicitly justified using the SMI. On the other hand, reviewers using the SMI may be overly lenient in their interpretation of the criteria. (This possibility did receive a preliminary, but intensive review in the validity investigation described in chapter IV.) The relative biases of the two instruments cannot be assessed without applying both to the same set of medical records and submitting discrepant evaluations for review by medical experts. Future independently-conducted studies of this type are called for in order to interpret the meaning of any differences in the magnitude of nonacute hospital utilization reported using the two instruments.

B. FACTORS ASSOCIATED WITH NONACUTE ADMISSIONS

This section and the following one present results of both the descriptive and multivariate analyses performed to identify factors associated with nonacute utilization. Since it was felt that the factors contributing to nonacute admissions might be very different from those contributing to nonacute days, two sets of analyses were performed: one for admissions and one for days.

These analyses were aimed at addressing the issues which had been raised in previous research on the subject of nonacute hospital utilization. Issues we sought to address specifically were:

- what kinds of patient stays are most susceptible to nonacute use, i.e.,
    - what kinds of patients
    - how long is the length of stay for these patients
    - in what part of the stay do the nonacute days occur

- how much nonacute utilization is associated with inefficient scheduling or poor discharge planning, e.g.,
    - Friday admissions for Monday surgery
    - discharge delayed over the weekend
    - other scheduling delays which occur in the treatment phase of the stay
    - other delays in discharge planning which occur at the end of the stay
- how much nonacute utilization is associated with differences in hospital structure, resource availability, delivery style, management, and financial incentives
- how much nonacute utilization is associated with reimbursement mechanisms (especially insurance coverage) which encourage overuse of hospital care
- how much nonacute utilization is due to the unavailability of subacute care
- how much nonacute utilization is associated with area factors which might create induced demand, e.g.,
    - the number of physicians per capita
    - hospital beds per capita
- how much of the regional variation in utilization rates (i.e., admission rates and ALOS) can be attributed to nonacute utilization

One research question conspicuously lacking from the above list concerns the degree to which nonacute utilization is related to variation in physician characteristics. The individual practice style of physicians has, in fact, been hypothesized to be by far the greatest factor in nonacute utilization, since it is the physician who is the decision-maker in most aspects of the hospital stay from admission to discharge. The sensitivity of the study precluded obtaining any information on individual physician characteristics from participating hospitals. It is likely, therefore, that a substantial amount of the variation in nonacute use will remain unaccounted for in the following analyses. We will attempt to infer whenever possible from the pattern of results what proportion of nonacute utilization is attributable to practices under the control of physicians.

1. Descriptive Analyses

In the descriptive analysis, the proportion of nonacute admissions was examined for various patient, hospital, and area characteristics. In addition, the reasons supplied by the reviewer as to the most likely cause of a nonacute admission were analyzed. The results are presented in Exhibits V-3, V-4, V-5, and V-6. Unless otherwise noted, the standard errors presented in the exhibits were calculated using the generalization function described in Chapter IV. As noted, the generalization function for the relvariance of estimated admissions is:

$$V^2_{x'} = .0001956 + \frac{11,643}{x'}$$

In applying this function to estimate the variance of any ratio presented in the following exhibits, the V^2 term is first found for the numerator of the ratio, then, using the same formula, the V^2 term is found for the denominator. The latter is then subtracted from the former to obtain the approximate relvariance of the ratio itself. This is multiplied by the square of the ratio to obtain the absolute rather than relative variance. The square root of this quantity is then taken to obtain the standard error of the ratio.

To test differences between two ratios in the table, the t-statistic is used as follows:

$$t = \frac{r_1 - r_2}{\sqrt{S^2_{r_1} + S^2_{r_2}}}$$

where

r_1 = the first ratio estimate
r_2 = the second ratio estimate
S_{r_1} = the standard error of the first ratio estimate
S_{r_2} = the standard error of the second ratio estimate

EXHIBIT V-3

PERCENTAGE OF NONACUTE ADMISSIONS BY PATIENT AND STAY FACTORS

	r 5.7%		(N_N/N_T)	s.e.
EXPECTED SOURCE OF PAYMENT				
Medicare	4.3%		(502/11,722)	0.64%
Medicaid	5.6%		(133/2,402)	1.61%
Other Government	8.3%	ns	(42/509)	4.19%
Worker's Compensation	8.8%		(57/650)	3.80%
Blue Cross	5.5%		(340/6,163)	0.98%
Commercial Insurance	6.0%		(531/8,789)	0.86%
HMO	0.0%		(0/185)	--
Self Pay Only	9.3%		(129/1,396)	2.66%
Other	12.8%		(85/670)	4.43%
Not Available	6.9%		(105/1,515)	2.21%
PRIMARY PLACEMENT				
Medical	9.1%	* * *	(1,495/16,351)	0.77%
Surgical	2.4%		(340/14,027)	0.44%
Psychiatric/Alcoholism	3.7%		(34/922)	2.12%
ICU/CCU/Burn Care or Trauma	1.2%		(27/2,265)	0.78%
Other	0.0%		(0/205)	--
Not Available	13.0%		(30/230)	7.55%
SOURCE OF ADMISSION (exc. Elec Surg)				
Emer. Rm., Working Week	5.2%		(292/5,608)	1.01%
Emer. Rm., Weekend or After Hours	6.7%	*	(94/1,401)	2.28%
Transferred from other unit or facility	1.8%	*	(19/1,073)	1.38%
Routine	9.7%		(1,226/12,603)	0.90%
Not Available	10.7%		(95/883)	3.55%
PATIENT AGE				
16-29	6.7%		(496/7,369)	1.00%
30-44	6.5%		(395/6,096)	1.08%
45-64	6.3%		(580/9,139)	0.86%
65-74	4.1%	*	(233/5,636)	0.90%
75+	3.9%		(221/5,631)	0.87%
Not Available	0.0%		(0/131)	--
PATIENT DISPOSITION				
Terminal Cancer Patient	17.1%		(48/282)	7.67%
Other Terminal Patients	0.0%		(0/809)	--
Patient Disch. to Subacute Care	6.7%	ns	(163/2,404)	1.73%
Other Patients	5.7%		(1,707/30,095)	0.46%
Not Available	1.8%		(7/411)	2.30%

* = significant at the .05 level

EXHIBIT V-3 (CONTINUED)

	r 5.7%		(N_N/N_T)	s.e.
MAJOR ICDA CLASS OF PRINCIPAL DIAGNOSIS				
Digestive	13.3%		(593/4,461)	1.74%
Mental	10.6%		(163/1,541)	2.68%
Musculoskeletal	9.6%	*	(181/1,893)	2.32%
Symptoms	9.6%		(180/1,874)	2.32%
Trauma	5.6%		(187/3,321)	1.36%
Respiratory System	4.3%		(126/2,927)	1.28%
Neoplasms	4.1%		(93/2,274)	0.98%
Other	3.9%		(105/2,719)	1.27%
Nervous System	3.7%	*	(61/1,658)	1.59%
Genito-Urinary	3.5%		(114/3,236)	1.10%
Pregnancy	2.4%		(60/2,505)	1.04%
Circulatory	1.1%		(62/5,593)	0.47%
LENGTH OF STAY (in days)				
1	4.2%	*	(135/3,198)	1.21%
2	5.8%		(262/4,529)	0.86%
3-5	7.9%		(872/11,078)	0.88%
6-15	5.0%	*	(585/11,829)	0.69%
16+	2.1%		(71/3,367)	0.84%
PHASE OF STAY UPON ADMISSION				
Work-Up	12.1%		(786/6,519)	1.38%
Pretreatment	2.5%		(201/8,106)	0.59%
Surgical Treatment	1.6%	* * *	(58/3,595)	0.71%
Medical Treatment	6.7%		(631/9,443)	0.88%
Not Available (Admission Day Not Abstracted)	4.0%		(250/6,339)	0.85%
PERCENTAGE OF STAY SPENT IN WORKUP				
0-25%	4.1%		(1,202/29,553)	0.40%
26-75%	9.6%	* *	(240/2,514)	2.01%
76-100%	25.2%		(483/1,919)	3.38%

r = ratio estimate of % nonacute admissions

N_N = estimated nonacute admissions

N_T = total estimated admissions

s.e. = standard error of r (estimated using the generalization function unless otherwise denoted by e, in which case standard error was calculated directly using the variances and covariances generated by the balanced repeated replication process).

ns = not significant

* = significant at the .05 level

NOTE: Brackets above denote that significance tests were conducted to compare the two ratios indicated by the bracket endpoints.

a. Patient Characteristics and the Rate of Nonacute Admissions

Exhibit V-3 presents the proportion of nonacute admissions for selected patient characteristics, including expected source of payment, primary placement, source of admission, patient age, patient disposition, the major ICDA class of the principal diagnosis, length of stay, phase of stay upon admission, and the percentage of the stay spent in a work-up phase. The major findings are outlined below.

i. Pay source

No significant differences were found among the four major pay sources (i.e., Medicare, Medicaid, Blue Cross, and other commercial insurance) with respect to the percentage of nonacute admissions. As can be seen in Exhibit V-3, these range from a low of 4.3% for Medicare patients to a high of 6.0% for patients covered by other commercial insurance. Although it appears that self paying patients, patients whose hospitalizations are covered by other government sources and patients covered by Worker's Compensation have higher than average rates of nonacute admissions, the standard errors for these groups are large due to the small sampled numbers. As a consequence, these differences are not statistically significant. The number of patients in the sample covered by an HMO is also very small so that the 0% rate of nonacute admissions found for this population cannot be generalized with any degree of confidence.

ii. Primary placement

The majority of patients in the sample were admitted to either a medical or surgical service, the former accounting for 48% and the latter for 41% of all admissions. As can be seen, the rate of nonacute admissions for medical cases is nearly four times higher than that for surgical cases (9.1% versus 2.4%), a difference which is statistically significant. It should be recalled, however, that the appropriateness of elective surgeries (i.e., those that did not otherwise meet a criterion) was determined only on the basis of whether the procedure was one that required an inpatient setting. The appropriateness or necessity of the surgery itself was not evaluated. Most admissions in which a major procedure was performed were evaluated as appropriate automatically, therefore, and the low nonacute admission rate for surgical services reflects this.

In addition to the difference between medical and surgical services, medical services have a much higher proportion of nonacute admissions than either psychiatric and alcoholism units (9.1% versus 3.7%) or ICU and other critical care units (9.1% versus 1.2%). Both these differences are significant.

iii. Source of admission

Preliminary analysis showed no significant differences between patients admitted routinely and those admitted through the emergency room. This was true both for patients admitted through the emergency room during normal working hours and those admitted on the weekend or after hours. Because elective surgery patients are often admitted through the emergency room on weekends or after hours, subsequent analyses excluded all patients admitted for elective procedures (defined as patients not meeting a criterion but appropriately hospitalized on the basis of a procedure which requires an acute care setting). As shown in Exhibit V-3, with the elimination of these admissions, a significant difference between emergency room admissions and routine admissions becomes evident: whereas less than 6% of patients admitted through an emergency room are inappropriately admitted, almost 10% of those admitted routinely are. This difference is statistically significant.

Routine admissions are also significantly more likely to be inappropriate than admissions for patients transferred or referred from another unit or facility, a difference of 9.7% versus 1.8%. The latter category includes patients transferred from another acute facility, a skilled nursing facility or an ambulatory surgery unit.

iv. Patient age

An examination of admissions by patient age shows a steady decrease in the proportion of nonacute admissions as patient age increases. Whereas 6.7% of patients in the 16 to 29 year old category are inappropriately admitted, only 3.9% of those 75 years of age or older are inappropriately admitted. This is most likely due to the increased severity of the health problems that occur among the elderly and to the fact that comorbid conditions associated with age will necessitate admissions for the elderly that would not be appropriate for younger patients.

v. Patient disposition

Logically, the manner in which a patient is discharged should have more effect on the frequency of nonacute days (particularly those at the end of the stay) than of nonacute admissions. For two classes of patients, however, the discharge status may be indicative of alternate care placement problems which could also cause inappropriate admissions. These are (1) patients who are discharged to subacute facilities and (2) terminally-ill patients who are hospitalized for humanitarian reasons. We, therefore, examined the nonacute admission rate for patients discharged to a subacute facility (SNF or ICF) and for patients who died while in the hospital. In examining the latter group, it was recognized that the class of patients who die in the hospital may be comprised not only of the terminally ill (who are more apt to be inappropriately hospitalized) but those who are in an acute or crisis condition (and are less apt to be inappropriately hospitalized). To distinguish these two types of patients, we defined as "terminally-ill" only those patients dying in the hospital who also had a primary diagnosis code for malignant neoplasms (ICD-9-CM diagnosis codes 140 to 208).

As can be seen in Exhibit V-3, patients discharged to subacute facilities have essentially the same rate of nonacute admissions as other patients (6.7% and 5.7%). They do not, therefore, appear to constitute a group which is particularly "prone" to inappropriate hospitalization. The percentage of nonacute admissions for terminal cancer patients, however, is, at 17.1%, substantially higher than average. In contrast, that for other (i.e., non-cancer) patients dying in the hospital is low (0%). Although these figures are suggestive, sample sizes for both the latter groups are small and the differences do not achieve statistical significance.

vi. Major ICDA class of principal diagnosis

As shown in Exhibit V-3, there are wide differences by body system in the rate of nonacute admissions. Digestive disorders (at 13.1%), mental disorders (10.6%), musculoskeletal problems (9.6%), and symptoms and ill-defined conditions (9.6%) fall above the mean in the proportion of nonacute admissions. All other categories fall below the mean. The category with the lowest proportion of nonacute admissions is circulatory disorders for which only 1.1% of the admissions were found to be nonacute.

vii. Length of stay

The rate of nonacute admissions demonstrates a curvilinear relationship with length of stay, with the highest rate being for stays of moderately short length, i.e., 3 to 5 days. Admissions for these stays are nonacute in almost 8% of the cases. In contrast, one day stays are inappropriate only 4% of the time, a difference which is statistically significant. The stays with the lowest percentage of nonacute admissions, however, are the longest stays, (i.e., of 16 days or more). Admissions for these stays were judged unnecessary only 2.1% of the time. This is also significantly different from the rate found for 3 to 5 day stays, but not significantly different from that for one day stays.

viii. Phase of stay upon admission

As part of the daily level of service evaluation, reviewers were required to indicate the phase of the stay. There were six phases in all, defined as follows:

- workup: includes days the patient is undergoing an admission work-up or awaiting physician or consultant visits prior to determination of therapeutic course. This may be because the etiology of the patient's complaint remains undetermined, tests are not complete, or observation is continuing prior to initiation of definitive therapy.
- pretreatment: used once a therapeutic course has been determined, but has not yet been started. This would include pre-operative preparation time. A patient can enter the hospital in phase 2 if, prior to admission, the course of therapy has been determined, such as scheduling of surgery.
- surgical treatment: surgical treatment phase in progress including the day of surgery and post-operative observation
- medical treatment: medical treatment in progress including drug therapy, respiratory therapy, physical therapy, psychotherapy or traction
- transition: defined as transition from one therapeutic regimen to another (as when complications develop or the course of treatment is re-evaluated)
- post-treatment: planned course of therapy has been completed

As shown in Exhibit V-3, there are wide differences in the rate of nonacute admissions depending upon the phase of treatment upon admission. The highest proportion of nonacute admissions is found for patients in a workup phase upon admission. These patients were judged to be inappropriately hospitalized 12.1% of the time. In contrast, those entering the hospital with the treatment course already planned (pretreatment phase) are inappropriately hospitalized only 2.5% of the time, and those entering directly into surgical treatment are inappropriately admitted only 1.6% of the time. (Again, it must be recalled that performance of major surgery was itself a criterion justifying admission to an acute care setting). Patients entering directly into a medical treatment phase are intermediate, being inappropriately hospitalized 6.7% of the time. Patients admitted into a workup phase, therefore, are inappropriately admitted twice as often as the next highest category, patients admitted into a medical treatment phase. This difference is statistically significant.

ix. Percentage of the stay spent in workup

Altogether, patients admitted in a workup phase represent approximately 24% of the admissions. (This percentage cannot be estimated exactly because information on the first day of the stay was not abstracted for some stays.) Of these patients, approximately 29% have stays primarily devoted to diagnostic workup, (i.e., over 75% of the days were spent in a workup phase). For another 39%, one-quarter to three-quarters of the stay is spent in workup.

As the percentage of the stay devoted to workup increases, so does the likelihood of an inappropriate admission. Exhibit V-3 presents the percentage of nonacute admissions by the percentage of the stay spent in workup. As can be seen, hospital stays which are primarily (i.e., over 75% of the stay) devoted to diagnostic workup are inappropriate 25.2% of the time. Those in which less than a quarter of the stay is devoted to workup are inappropriate only 4.1% of the time. Those which are intermediate in terms of the percentage of the stay devoted to workup (i.e., between one-quarter and three-quarters of the stay) are also intermediate, at 9.6%, in terms of the proportion of inappropriate admissions.

b. Hospital Characteristics and the Rate of Nonacute Admissions

Exhibit V-4 presents the distribution of nonacute admissions by selected hospital characteristics, including bedsize, ownership, teaching status, occupancy rate, and expenditures per bed. These are discussed in turn below.

i. Bedsize

As shown in Exhibit V-4, smaller hospitals are more likely to have nonacute admissions than larger hospitals. The rate of nonacute admissions is 7% for hospitals with fewer than 99 beds, as well as for those having 100-299 beds. Only slightly more than 4% of the admissions for larger hospitals (both those with 300 to 499 beds and those with over 500) beds are inappropriate. The difference between these two sets of hospitals is statistically significant.

ii. Ownership

Although it appears that the rate of nonacute admissions is higher for proprietary hospitals (7.4%) than nonprofit and nonfederal government hospitals (5.7% and 4.8% respectively), this difference is not significant.

EXHIBIT V-4

PERCENTAGE OF NONACUTE ADMISSIONS BY HOSPITAL CHARACTERISTICS

	r 5.7%		(N_N/N_T)	s.e.	
BEDSIZE					
6-99	7.0%		(422/6,056)	1.12%	
100-299	7.0%	*	(808/11,612	0.81%	
300-499	4.2%		(359/8,511)	0.74%	
500+	4.3%		(336/7,822)	0.78%	
OWNERSHIP					
Government, nonfederal	4.8%		(380/7,944)	0.82%	0.65%e
Nonprofit	5.7%	ns	(1,288/22,576)	0.53%	
Proprietary	7.4%		(257/3,481)	1.64% e	
TEACHING STATUS					
Medical School Affiliated	3.7%		(379/10,118)	0.64%	
Residency Program Only	5.4%	*	(130/2,399)	1.57%	
Nonteaching	6.6%		(1,417/21,485)	0.58%	
OCCUPANCY RATE					
0-59%	7.9%		(317/4,035)	1.62% e	
60-79%	6.9%	*	(1,061/15,435)	0.70%	
80-89%	4.1%		(478/11,607)	0.63%	
90%+	2.4%		(69/2,924)	0.80% e	
HOSPITAL EXPENDITURES PER BED					
0-$54,807	7.2%		(606/8,394)	0.96%	
$54,808-69,696	6.9%		(645/9,322)	0.89%	
$69,697-85,231	4.8%	*	(373/7,860)	0.83%	
>$85,231	3.6%		(301/8,425)	0.70%	

r = ratio estimate of % nonacute admissions

N_N = estimated nonacute admissions

N_T = total estimated admissions

s.e. = standard error of r (Estimated using the generalization function unless otherwise denoted by e, in which case standard error was calculated directly using the variances and covariances generated by the balanced repeated replication process.)

ns = not significant

* = significant at the .05 level

NOTE: Brackets above denote that significance tests were conducted to compare the two ratios indicated by the bracket endpoints.

iii. Teaching status

Hospitals with a medical school affiliation have a significantly lower rate of nonacute admissions (3.7%) than nonteaching hospitals (6.6%). This difference is significant. Hospitals with residency programs (but not a medical school affiliation) fall in between these extremes with a rate of 5.4%.

iv. Occupancy rate

A direct relationship was found between a hospital's occupancy rate and the proportion of nonacute admissions. In hospitals for which the occupancy rate is less than 60%, almost 8% of the admissions are nonacute. This rate declines steadily as the occupancy rate increases. For hospitals with an occupancy rate greater than or equal to 90%, the proportion of nonacute admissions is slightly over 2%. The difference between these two extremes is statistically significant.

v. Hospital expenditures per bed

The rate of nonacute admissions also decreases steadily as the expenditures per bed increase. For hospitals with expenditures greater than $85,231 per bed, only 3.6% of the admissions are nonacute. (These hospitals accounted for approximately 25% of all admissions.) For hospitals spending less than $54,808 per bed (again accounting for about 25% of the admissions), 7.2% of the admissions are nonacute.

c. Area Characteristics and the Rate of Nonacute Admissions

Exhibit V-5 presents the proportion of nonacute admissions by selected area characteristics, including region, SMSA size, nursing home beds per capita, hospital beds per capita, physicians per capita, and the percentage of the population below the poverty line. These are discussed below.

EXHIBIT V-5

PERCENTAGE OF NONACUTE ADMISSIONS BY AREA CHARACTERISTICS

	r 5.7%		(N_N/N_T)	s.e.
REGION				
South	7.9%		(466/9,847)	0.86%
West	5.0%		(264/6,870)	1.00%
North Central	4.7%	*	(906/11,520)	0.73%
Northeast	3.9%		(289/5,765)	0.80%
SMSA SIZE				
Non-SMSA	8.0%		(746/9,305)	0.96 %
SMSA of <100K to 1 Mil	4.3%	* *	(468/10,830)	0.66 %
SMSA > 1 Mil	5.1%		(711/13,866)	0.64 %
NURSING HOME BEDS PER CAPITA (x1,000)				
0-4.0	6.8%		(704/10,290)	0.84 %
4.1-7.0	5.6%	*	(830/14,767)	0.64 %
>7.0	4.4%		(391/8,943)	0.74 %
HOSPITAL BEDS PER CAPITA (x1,000)				
0.3-4.0	5.4%		(550/10,193)	0.76 %
4.0-6.0	5.9%	ns	(803/13,596)	0.69 %
6.0 +	5.6%		(572/10,212)	0.78 %
PHYSICIANS PER CAPITA (x1,000)				
0.1 to 1.2	7.2%	*	(912/12,689)	0.78%
1.2 to 2.0	4.8%		(564/11,860)	0.67%
2.0 +	4.8%		(450/9,452)	0.75%
PERCENTAGE OF POPULATION BELOW THE POVERTY LEVEL				
0-7.5	4.5%		(439/9,721)	0.72%
7.5-10.0	3.9%	*	(279/7,184)	0.78%
10.0-15.0	7.7%		(807/10,494)	0.89%
15.0+	6.1%		(401/6,602)	1.01%

r = ratio estimate of % nonacute admissions

N_N = estimated nonacute admissions

N_T = total estimated admissions

s.e. = standard error of r (estimated using the generalization function unless otherwise denoted by e, in which case standard error was calculated directly using the variances and covariances generated by the balanced repeated replication process).

ns = not significant

* = significant at the .05 level

NOTE: Brackets above denote that significance tests were conducted to compare the two ratios indicated by the bracket endpoints.

i. Region

Wide variation by region is evident in Exhibit V-5, with almost 8% of the admissions in the South being nonacute, compared with 4% in the Northeast. This difference is statistically significant. The variation in the percentage of admissions by region is, in fact, similar to the frequently observed regional variation in admission rates. Notably, the percentage of nonacute admissions is highest in the South where the admission rate per 1000 is also highest and lowest in the Northeast where the admission rate is also the lowest. The proportion of nonacute admissions for the West found in this study, however, is actually somewhat higher than that for the North Central region even though the latter has a higher admission rate. This difference is not statistically significant. These findings, which are apparent at a descriptive level, will be altered somewhat in the multivariate analyses, described later, in which many of the differences among regions are controlled.

ii. Other area-level differences

In addition to the regional variation in nonacute admission rates, there are also significant differences in the proportion of nonacute admissions by SMSA size, nursing home beds per capita, physicians per capita, and the percentage of the population below the poverty level. In particular, non-SMSAs tend to have a higher rate of nonacute admissions (8%) than either SMSAs of over 1,000,000 (5%) or small-to moderate-size SMSAs (4%). Areas with more nursing home beds per capita and more physicians per capita have a lower rate of nonacute admissions, the latter trend running counter to the hypothesis that admissions increase unnecessarily as the physician to population ratio increases. Areas which have over 10% of their population living below the poverty level also have a higher proportion of nonacute admissions (7.1%) than those with less than 10% (4.2%). No significant relationship is apparent between hospital beds per capita and the rate of nonacute admissions.

d. Reviewer-supplied Reasons for Nonacute Utilization

For each admission judged to be nonacute, reviewers were asked to provide the most likely reason for the occurrence of that admission. As part of this process, reviewers also indicated whether the patient required subacute care or no institutional care at all. The reasons given by reviewers for nonacute admissions are summarized in Exhibit V-6. Reviewers were not asked to distinguish patient or family responsibility from physician responsibility in assigning reasons for a nonacute admission or day. It was felt that this information would not be reliably indicated in the medical record and, furthermore -- given the complexity of patient-physician interactions -- such attributions are not easily made under any circumstances. The reason lists used by abstractors can be found in Exhibit III-5 in Chapter III.

Of the patients who were inappropriately hospitalized, reviewers felt that 13% required institutional care at a subacute level and 87% did not require institutional care of any kind. Of the 87% who required no institutional care, 3% were terminal patients whom reviewers felt had been hospitalized for humanitarian reasons. Another 7% were those for whom no home care was available. An additional 36% were felt by reviewers to have been hospitalized at the insistence of the patient, family or physician. In 41% of the cases, however, reviewers were unable to deduce the reason the patient had been inappropriately hospitalized. A chi-square test was performed to compare nonacute admissions in this latter category (i.e., those for whom no reason could be deduced) with those attributed to patient or physician insistence. Admissions were compared with respect to rural/urban location (defined as non-SMSAs versus SMSAs), patient age (over 75 years versus under 75 years), and pay source (Medicaid, no charge, self-pay, other government, versus all others). No significant differences were found between the two categories of reasons.

Among the 13% of patients requiring lower level care, reviewers felt that 4.5% had been hospitalized because alternate care was unavailable or because of other placement difficulties. Another 0.5% were terminal patients hospitalized for humanitarian reasons. Almost 5%, however, were felt to have been hospitalized because of patient, family, or physician insistence. A final 3% were patients for whom reviewers could not deduce the reason for the inappropriate admission. Therefore, of all patients requiring subacute care, only 5% to 8% (counting the 3% for whom no reason could be deduced) can be considered inappropriately hospitalized because of the unavailability of alternate care facilities. If the 7% of patients who require home care but have none available are also viewed as an alternate care problem, along with the 3% of terminal patients who do not require institutional care but have been

EXHIBIT V-6

REVIEWER REASONS FOR NONACUTE ADMISSIONS
(Includes Only Admissions Judged Nonacute by Reviewer)

	r	N
PATIENT REQUIRES NO INSTITUTIONAL CARE		
Terminal Patient for Humanitarian Reasons(11)[1]	3.1%	(38,340)
Patient, Family, Physician Insist (12)	36.0%	(444,836)
No Home or Alternate Care Available (15)	7.1%	(87,085)
No Reason Can Be Deduced (16,17)	41.1%	(507,685)
TOTAL	87.3%	(1,077,946)
PATIENT REQUIRES SUBACUTE CARE		
Terminal Patient for Humanitarian Reasons(21)	0.5%	(5,777)
Lower Care Unavailable/Difficulty Placing(23,24)	4.5%	(55,791)
Patient, Family, Physician Insist(26)	4.7%	(57,596)
No Reason Can Be Deduced (27,28)	3.1%	(37,851)
SUB-TOTAL	12.7%	(157,015)
TOTAL	100.0%	(1,234,961)

r = ratio estimate of % reasons in each category over total admissions judged nonacute.

N = estimated number of reasons in each category (total = 1,234,961) hospitalized for humanitarian reasons, then from 15% to 18% of the inappropriate admissions were judged by reviewers to be attributable to the unavailability of alternate care.

---

[1] Reason code numbers from the original list are supplied in parentheses.

[2] Column total is less than the 1,925,000 total estimated nonacute admissions due to missing reason codes for evaluations that were over-ridden following data collection.

2. Multivariate Analyses

Two types of analyses were performed in the multivariate examination of factors associated with nonacute utilization. The first was at the patient level and the second at the hospital level. These are discussed in turn below.

a. Patient-level Model

Since the appropriateness of an admission is a binary value (1 for nonacute, 0 for acute) the logit procedure was used in performing regression analyses for predicting nonacute admissions. The procedure used was that supported by the Statistical Analysis System (SAS) Software Package. In performing the analysis, cases were weighted by the original sampling and nonresponse weights, i.e., those calculated prior to application of the first-stage and second-stage ratio estimators. The latter are useful primarily in improving the precision of ratio estimates, and have no inherent advantage when the dependent measure is a binary variable.

Three sets of variables were used in developing the model. These were (1) patient and patient stay characteristics, (2) hospital characteristics, and (3) area characteristics. The variables used in the analysis are presented in Exhibit V-7. The patient characteristics selected for analysis are generally those which had been identified in the descriptive analysis as having the strongest relationships with nonacute admissions. These were recoded as necessary into dummy variables for the regression model.

The hospital and area characteristics used in performing the logit regression were also those used in the descriptive analyses. In developing the hospital-level model described later in this chapter, a great many hospital and area characteristics were examined as potential candidates for that model. The eighteen variables selected for that model were the most theoretically relevant among these. In performing the patient-level logit analyses we sought to

EXHIBIT V-7

VARIABLES USED IN THE PATIENT-LEVEL MODELS

DEPENDENT VARIABLES:

ADMINAP: Coded 1 if the admission is nonacute. Coded 0 otherwise.

DAYINAP: Coded 1 if the day is nonacute. Coded 0 otherwise.

PATIENT/STAY FACTORS:

LTCRISK: Coded 1 if the patient is greater than 74 years old and discharged to subacute care (skilled nursing facility, other subacute facility, or organized home care program). Coded 0 otherwise.

CAREPAY: Coded 1 if the primary paysource is Medicare. Coded 0 otherwise.

CAIDPAY: Coded 1 if the primary paysource is Medicaid. Coded 0 otherwise.

PRIPAY: Coded 1 if the primary paysource is Blue Cross or other commercial insurance. Coded 0 otherwise.

SPECPLAC: Coded 1 if the primary placement unit is an ICU, CCU, Burn Care Unit or Shock/Trauma unit. Coded 0 otherwise.

URGADMIT: Coded 1 if the patient is admitted through the Emergency Room and is not admitted for elective surgery. Coded 0 otherwise.

P_WORKUP: A continuous variable equal to the percentage of the stay that was spent in a workup phase.

PATAGE: A continuous variable equal to the patient's age in years.

SHORTLOS: Coded 1 if the length of stay is one or two days. Coded 0 otherwise.

MEDLOS: Coded 1 if the length of stay is three, four or five days. Coded 0 otherwise.

LONGLOS: Coded 1 if the length of stay is over 30 days. Coded 0 otherwise.

MENTAL: Coded 1 if the primary diagnosis is in the ICDA Class for Mental Disorders (ICD-9-CM codes 290-319). Coded 0 otherwise.

SYMPTOMS: Coded 1 if the primary diagnosis is in the ICDA Class for Symptoms, Signs, and Ill-Defined Conditions (ICD-9-CM codes 780-799). Coded 0 otherwise.

DIGPROB: Coded 1 if the primary diagnosis is in the ICDA Class for Digestive Disorders (ICD-9-CM codes 520-579). Coded 0 otherwise.

MUSKEL: Coded 1 if the primary diagnosis is in the ICDA Class for Muskuloskeletal Problems (ICD-9-CM codes 710-739). Coded 0 otherwise.

PROC: Coded 1 if a UHDDS class 1 procedure is performed. Coded 0 otherwise.

WORKUP: Coded 1 if a day is in a workup phase. Coded 0 otherwise. Used in day level analyses only.

ENDSTAY: Coded 1 if a day is in the final portion of the stay. Coded 0 otherwise (used in day level analyses only).

SNFDISP: Coded 1 if the patient is discharged to a skilled nursing facility. Coded 0 otherwise.

HOSPITAL CHARACTERISTICS:

TOTLBEDS: A continuous variable equal to the bedsize of the hospital.

PROFIT: Coded 1 if the hospital is proprietary. Coded 0 for nonprofit and government, nonfederal hospitals.

OCCRATE: A continuous variable equal to the occupancy rate of the hospital.

PROFIT*OCC: An interaction term equal to PROFIT * OCCRATE

EXPBED: A continuous variable equal to total hospital expenditures in 1980 divided by hospital bedsize.

TEACH: Coded 1 if a hospital is affiliated with a medical school or has a residency program. Coded 0 otherwise.

AREA CHARACTERISTICS:

N-EAST: Coded 1 if the hospital is located in the Northeast region. Coded 0 otherwise.

SOUTH: Coded 1 if the hospital is located in the Southern region. Coded 0 otherwise.

WEST: Coded 1 if the hospital is located in the Western region. Coded 0 otherwise.

RURAL: Coded 1 if the hospital is located in a non-SMSA. Coded 0 otherwise

HBEDPERK: A continuous variable equal to the number of short-term general hospital beds per capita in 1981 in the FIPS county area.

MDSPERK: A continuous variable equal to the number of physicians in 1979 (direct patient care only) per capita in the FIPS county area.

NBEDPERK: A continuous variable equal to the number of nursing home beds per capita in 1978 in the FIPS county area.

POVERTY: A continuous variable equal to the percentage of the population in 1970 in the FIPS county area with incomes under the poverty level.

minimize the number of hospital and area level variables that were examined in addition to the patient characteristics. The number of variables was, therefore, reduced from 18 to 14 by exclusion of those with the weakest hypothesized relationships with nonacute utilization.

The first step in the analysis was to perform a stepwise regression for each set of characteristics. Exhibit V-8A presents the stepwise regression analysis for the set of patient variables. Eight of the 14 patient variables contribute significantly to the equation[1]. That with the largest effect is P_WORKUP, or the percentage of days in the stay spent in a workup phase. This is positively related to the probability of a nonacute admission. Admissions for four diagnostic categories also have a high probability of being a nonacute. These are digestive problems (DIGPROB), which has the largest effect, followed by mental disorders (MENTAL), musculoskeletal problems (MUSKEL), and symptoms and ill-defined conditions (SYMPTOMS). Admissions for patients placed in a critical care unit (SPECPLAC), such as an ICU, CCU, burn, shock or trauma unit, have a lower probability of being nonacute, as do non-elective admissions through the emergency room (URGADMIT). Patient age (PATAGE) is negatively associated with the likelihood of a nonacute admission. The effect of the intercept is also significant, indicating that the probability of having a nonacute admission is significantly greater than zero. The model is significant (Chi-square = 238.9, $p < .0001$) and the rank correlation between the predicted probability and the actual value is .48.

The variables which are significant in the model described above are the same as those found significant in the descriptive analyses, with two exceptions. There is no significant effect in the model of MEDLOS (representing stays of three to five days), although a weak effect is evident prior to entry of any of the variables into the equation (as shown in the left-hand column). Its lack of significance is most likely due to the fact that, as a dummy variable, three to five day stays are contrasted to all other stays; whereas in the descriptive analyses, they were contrasted to the two extremes in terms of length of stay. In addition, although Medicare pay source (CAREPAY) was not

---

[1] In this, and all subsequent exhibits, all effects significant at a .05 level are reported. Although many effects achieved the .01 and .001 levels of significance, these are not individually identified.

EXHIBIT V-8A

PATIENT-LEVEL ADMISSION ANALYSIS: PATIENT CHARACTERISTICS
STEPWISE LOGIT REGRESSION MODEL

All Variables Prior to Entry into Model:			Variables Entered Through Stepwise Regression:		
VARIABLE	BETA	CHI-SQUARE	VARIABLE	BETA	CHI-SQUARE
LTCRISK	0	0.92	Intercept	-3.11	290.65*
CAREPAY	-	5.58*	P-WORKUP	3.20	121.43*
PRIPAY	0	0.88	DIGPROB	1.21	62.02*
SPECPLAC	-	16.58*	SPECPLAC	-2.07	17.98*
URGADMIT	-	7.30*	MENTAL	1.73	29.98*
P_WORKUP	+	204.61*	MUSKEL	0.69	9.28*
SHORTLOS	0	0.00	SYMPTOMS	0.77	10.17*
MEDLOS	+	2.66	URGADMIT	-0.66	7.48*
LONGLOS	-	3.82	PATAGE	-0.01	7.19*
MENTAL	+	2.88			
SYMPTOMS	+	10.46*			
DIGPROB	+	74.53*			
MUSKEL	+	4.38*			
PATAGE	-	4.43*			

Model Chi-Square = 238.9*

Rank correlation between predicted probability and response = .48

* = significant at the .05 level.

significantly related to nonacute admissions in the descriptive analyses, it does appear to have a potentially significant effect in the logit model. Although it does not make a significant contribution to the final equation, its effect is significant prior to entry of any of the variables into the equation (as shown in the left hand column). Its failure to load in the equation is most likely due to a collinear relationship with patient age, the latter of which is a significant factor in the model.

Exhibit V-8B presents the stepwise logit regression equation developed for the hospital-level variables. Five variables were examined: bedsize (TOTLBEDS), profit versus nonprofit ownership status (PROFIT), occupancy rate (OCCRATE), a profit times occupancy rate interaction term (PROFIT*OCC), and annual hospital expenditures per bed (EXPBED). As can be seen in the left-hand column, three variables are significantly related to the probability of a nonacute admission prior to entry of any variables into the model. These are the same effects found significant in the descriptive analyses: namely, hospitals with more beds, a higher occupancy rate, and higher expenditures per bed have a lower probability of nonacute admissions.

In the final model, however, (shown in the right hand column), only expenditures per bed (EXPBED) has a significant effect, indicating that occupancy rate and bedsize are substantially correlated with expenditures per bed and contribute nothing above and beyond it in predicting nonacute admissions. This model is also significant (Chi-square = 22.82, $p < .01$) with a rank correlation coefficient between the predicted probability and actual value of .16.

Exhibit V-8C presents the stepwise regression model for area characteristics. Prior to entry into the model, five variables significantly predict the probability of a nonacute admission. These are similar to those found significant in the descriptive analyses. Namely, there is a higher probability of a nonacute admission in the South (SOUTH) and a lower probability in the Northeast (N_EAST) and West (WEST). Rural areas (RURAL) are more likely to have nonacute admissions, and areas with more physicians per capita (MDSPERK)

EXHIBIT V-8B

PATIENT-LEVEL ADMISSION ANALYSIS: HOSPITAL CHARACTERISTICS STEPWISE LOGIT REGRESSION MODEL

All Variables Prior to Entry into Model:			Variables Entered Through Stepwise Regression:		
VARIABLE	BETA	CHI-SQUARE	VARIABLE	BETA	CHI-SQUARE
TOTLBEDS	-	13.86*	Intercept	-2.13	142.22*
PROFIT	0	0.19	EXPBED	-1.2E-5	20.95*
OCCRATE	-	6.76*			
PROFIT*OCC	0	0.17			
EXPBED	-	20.92*			

Model Chi-Square = 22.82*

Rank Correlation between Predicted Probability and Response = .16

* = significant at the .05 level

EXHIBIT V-8C

PATIENT-LEVEL ADMISSION ANALYSIS: AREA CHARACTERISTICS STEPWISE LOGIT REGRESSION MODEL

All Variables Prior to Entry into Model:			Variables Entered Through Stepwise Regression:		
VARIABLE	BETA	CHI-SQUARE	VARIABLE	BETA	CHI-SQUARE
N_EAST	-	7.28*	Intercept	-3.05	408.8*
SOUTH	+	25.32*	SOUTH	0.50	14.24*
WEST	-	11.05*	RURAL	0.40	9.30*
RURAL	+	16.12*	HBEDPERK	-0.06	5.72*
HBEDPERK	0	1.50			
MDSPERK	-	10.42*			
NBEDPERK	0	1.32			
POPOLD	0	0.70			

Model Chi-Square = 40.97*

Rank Correlation between Predicted Probability and Response = .22

* = significant at the .05 level

are less likely to have nonacute admissions. Unlike the relationships found in the descriptive analysis, nursing home beds per capita (NBEDPERK) and poverty level (POVERTY) do not have a significant effect.

Only three of these variables contribute significantly to the final model. SOUTH and RURAL are both related to the probability of a nonacute admission in the predicted direction. WEST and N_EAST fail to enter the equation, although the effect for WEST borders on significance. MDSPERK, although significant at a bivariate level, fails to enter the equation, but HBEDPERK, although not significant at a bivariate level, is significantly related to nonacute use in the context of the other variables. The model is significant (Chi-square = 40.97, $p < .0001$) and the rank correlation between the predicted and actual value is .22.

Exhibit V-9 presents a model based on the combination of all patient, hospital, and area variables. The model is highly significant (Chi-square = 362.28, $p < .0001$) and the rank correlation between the predicted probability and the actual value is .55. In this model, all the patient characteristics which were significant in the stepwise patient-level model (Exhibit V-8A) remain significant, with the exception of patient age. These are, in addition, the largest effects in the model. No hospital-level effects are significant in this equation, most likely due to collinearity with the area characteristics. Of the latter, there remains a strong effect of region. The positive association between SOUTH and the probability of a nonacute admission is no longer significant. However, the negative associations with N_EAST and WEST have become so again. The negative association with hospital beds per capita (HBEDPERK) also remains significant.

b. Hospital-level Model

In addition to the models developed at the patient level, the percentage of nonacute admissions was aggregated at the hospital level to examine further the hospital and area factors associated with the percentage of nonacute admissions. In general, one might expect similar results between the patient level and hospital level analyses, unless the types of patients served by hospitals vary

EXHIBIT V-9

PATIENT-LEVEL ADMISSION ANALYSIS: PATIENT, HOSPITAL, AND AREA CHARACTERISTICS LOGIT REGRESSION MODEL

VARIABLE	BETA	CHI-SQUARE
Intercept	-2.22	12.68*
LTCRISK	0.06	0.10
CAREPAY	-0.34	1.81
PRIPAY	-0.31	2.91
SPECPLAC	-1.90	15.42*
URGADMIT	-0.69	8.09*
DIGPROB	1.18	55.96*
MUSKEL	0.72	9.50*
MENTAL	1.56	23.70*
P_WORKUP	3.18	107.23*
SHORTLOS	0.21	1.53
MEDLOS	0.08	0.24
LONGLOS	-0.84	1.30
PATAGE	-0.01	1.58
SYMPTOMS	0.78	10.08*
TOTLBEDS	-0.00	3.41
PROFIT	-0.44	0.28
OCCRATE	-0.00	0.47
PROFIT*OCC	-0.00	0.03
EXPBED	-0.00	0.19
N_EAST	-0.50	4.07*
SOUTH	0.37	3.69
WEST	-0.58	4.70*
RURAL	0.09	0.22
HBEDPERK	-0.08	4.37*
MDSPERK	0.12	1.54
NBEDPERK	-0.03	1.15
POVERTY	0.07	2.71

Model Chi-Square = 362.28*

Rank Correlation between predicted probability and response = .55

* = significant at the .05 level

widely and serve to moderate the effects of the hospital and area level variables.

To perform the analysis, the percentage of nonacute admissions was first calculated for each hospital. No weight was used in calculating this percentage; since both the numerator and denominator are weighted by the same factor in calculating the percentage of nonacute admissions, the weight, in effect, cancels out. In performing the regression, however, each hospital was weighted by its stratum measure of size (i.e., the number of admissions in the stratum from which it was drawn) so that it represented the appropriate number of admissions in the universe.

Hospital variables were selected from the 1980 American Hospital Association Annual Survey of Hospitals which contains over 400 characteristics covering staffing, expenditures, facilities and utilization within the hospital. Area variables were selected from the Area Resource File (HRA, 1980), which provides data by FIPS county area on the health-related characteristics of the area, including facilities, personnel and population. Data for the year closest to the year of the study (i.e., 1981) were taken from this file.

A number of conceptual dimensions were considered in defining the parameters of the model. These were as follows:

Hospital Characteristics

(1) Size: Hospital size correlates with a number of differences in patient type and practice style. These differences may, in turn, translate into treatment patterns that affect the degree of nonacute utilization. Although the direction of the effect was not predicted in advance, the potential influence of this variable was felt to be such as to warrant investigation.

(2) Resource Intensity: This factor was defined as the variety and technological sophistication of services provided by the hospital. Again, the types of patients served and practice styles of hospitals differing in resource intensity are themselves quite different. These differences may

serve to promote or inhibit nonacute utilization. As with hospital size, there was no hypothesis concerning the direction of the effect of resource intensity with respect to the nonacute utilization. Although resource intensity may be associated with the overutilization of ancillary services, there is no particular reason to believe that it is associated with overutilization of the acute care setting itself.

(3) General Resource Availability: Independent of the concept of resource intensity is that of resource availability. In particular, it was hypothesized that hospitals with more resources -- including more RNs per bed (and fewer RN vacancies per bed), more technicians per bed, the availability of weekend laboratory and radiology staffing, organized rehabilitation and psychiatric treatment units -- might be more efficient and, therefore, less prone to unnecessary utilization than hospitals without these resources.

(4) Utilization-Control Resources: In addition to the general availability of resources, certain services might directly promote the efficient discharge of patients. Such services include the availability of an organized home care department, an organized discharge planning department, an ambulatory surgery unit, an associated long term care unit, and so forth.

(5) Practice Style: Although teaching hospitals tend to be larger and more resource intensive, it was hypothesized that they might have a unique practice style of their own that affects nonacute utilization. In specific, it has been hypothesized that teaching hospitals may tend to be more inefficient because they prolong stays for teaching purposes. Teaching hospitals also have greater resource availability, however, a factor which might promote efficiency. In the model, we wished to examine the effect of teaching status above and beyond the effects of size, resource intensity and resource availability, so teaching status was entered in the model after those representing the latter dimensions.

(6) Financial Organization and Incentives: The manner in which hospitals are organized and managed and their financial goals may affect their tendency to value or disvalue efficiency or even to consciously encourage or discourage overutilization. We were particularly interested in examining such variables as ownership, membership in a corporation, whether the hospital was run by a management organization, and so forth. Also included in this category was occupancy rate. With respect to the latter, it was felt that occupancy rate could have a direct effect on the tendency to unnecessarily admit patients or prolong their stays. It was also hypothesized that occupancy rate might interact with organizational structure in its effect. For example, profit hospitals with low occupancy rates might have an economic incentive to make up this deficit through unnecessary hospital utilization, whereas nonprofit hospitals might not.

Area Characteristics:

(1) Alternate Care Resource Availability: Considering that the unavailability of alternate care facilities might make a significant contribution to the problem of nonacute hospital utilization, this was considered a key dimension of the model. Variables examined for inclusion in the model were nursing home beds per capita, nursing home beds per aged population only, total long-term care hospital beds in the area and so forth.

(2) Supply-Induced Demand: The fact that physician-supply and hospital-bed supply per capita are highly correlated with utilization rates implies that some of this utilization is supplier-induced and, therefore, unnecessary. The model, therefore, incorporated these parameters to examine their relationship with the rates of unnecessary utilization measured in the study. Since hospital bed supply and physician supply are correlated, the physician effect was introduced first in constructing the admission model. This is because physicians were hypothesized to be more likely to control the admission rate. An effect of hospital-induced demand was, therefore, only of interest if it contributed above and beyond the effect of physician-induced demand. The converse was true of the day-level model, in which hospitals were considered more likely to influence the length of stay for financial reasons than physicians. Another variable

considered to relate to this dimension particularly in the admission analysis was the prevailing charge rate for physicians for the provision of hospital services. It was hypothesized that the higher the reimbursement to the physician for hospital services, the more incentive to hospitalize unnecessarily. Since the prevailing charge rate is correlated with physicians per capita, it was introduced into the model following that variable to identify whether it had an independent effect of its own.

(3) Population-Needs: The population needs with respect to health care vary across areas and may put different demands on hospitals. With older and/or sicker populations, hospitals might be used more often for needed care. In areas with older populations, however, there may also be more problems stemming from the lack of availability of alternate care, thereby putting more pressure on hospitals to provide subacute rather than acute care. In areas with higher poverty levels, hospitals may similarly be used as a mechanism for providing social services that are not available elsewhere rather than for acute care exclusively.

(4) Practice Style: Observed differences between areas in utilization rates and, thus, in nonacute utilization may be attributable to differences in practice style between physicians in the areas. Although a difficult dimension to operationalize, the proportion of older physicians in an area and the ratio of generalists to specialists were examined for inclusion in the model. Since these variables may be correlated with other area characteristics such as physicians per capita, they were introduced in the model after such variables in order to assess their unique contribution.

(5) Aggregate Characteristics: Both region and urban/rural location are associated with wide differences in per capita utilization rates. These differences may merely be the sum of the individual effects examined above or may instead (or additionally) reflect differences in practice style and philosophy unique to the area. Accordingly, these variables were examined last in the model to assess their contribution above and beyond the effects of the component characteristics.

In attempting to operationalize each dimension, a number of hospital and area characteristics were examined prior to specifying the final variables for the model. The final hospital level and area level variables were carefully selected since, with 253 hospitals in the sample, the degrees of freedom were limited. Selecting variables with no missing data was also considered essential. Since each hospital represents in essence a non-overlapping segment of the universe, the effect of dropping a hospital from the analysis creates problems in generalizing to the population as a whole.

To select the variables for the model, several correlation matrices were calculated among sets of variables, the first, measuring area level characteristics, the second, measuring hospital characteristics related to organization, structure, and financial characteristics and the third, measuring hospital characteristics related to resource availability, size, and complexity. These matrices are presented in Appendix L.

From among these variables, 18 were selected as being the most theoretically relevant and conceptually distinct. It should be noted that it was not possible to find distinct variables measuring Resource Intensity, General Resource Availability, and Utilization-Control Resources. Although there were a number of variables available to represent the latter dimension, (e.g., the availability of a home care department, the availability of a subacute unit, weekend laboratory and radiology department staffing, the presence of an organized discharge planning department, the number of nurses per bed, and so forth) all were highly correlated with hospital size and complexity, the latter measured simply by the number of separate services provided by the hospital. In addition, most of these variables -- including the number of services provided by the hospital -- had substantial missing data rates. Therefore, total hospital expenditures per bed (EXPED), which was also highly correlated with these resource variables, was used to represent all three dimensions.

The variables selected for the final model and the dimensions they represent are shown below. These are defined in Exhibit V-10A. (As shown, three regional dummy variables were created, one for the South, one for the West, and one for the Northeast. The North Central region was used as the contrast since nationally, it is intermediate in terms of both the admission rate and ALOS.)

Size:	TOTLBEDS
Resource Intensity General Resource Availability: Utilization-Control Resources	EXPBED
Practice Style:	TEACH
Financial Structure and Incentives:	PROFIT OCCRATE PROFIT*OCC
Supplier-Induced Demand:	MDSPERK HBEDPERK PREV_CHG
Alternate Care Availability:	NBEDPERK
Population-Needs:	POVERTY POPOLD NBEDOLD
Practice Style:	OLDPHY
Aggregate Characteristics:	RURAL N_EAST SOUTH WEST

The intercorrelations among these variables are presented in Exhibit V-10B. As can be seen, there are substantial correlations among a number of the variables. Bedsize has large positive correlations with expenditures per bed (r = .48), teaching status (r = .67), occupancy rate (r = .51), hospital beds per capita (r = .27) physicians per capita (r = .48), and prevailing charges (r = .38), and a negative relationship with rural location (r = -.44). Expenditures per bed has large correlations with the same variables in the same

EXHIBIT V-10A

HOSPITAL AND AREA VARIABLES SELECTED FOR THE HOSPITAL-LEVEL MODEL

DEPENDENT VARIABLES:

PCTIADM:	A continuous variable equal to the percentage of nonacute admissions calculated for each hospital.
PCTIDAY:	A continuous variable equal to the percentage of nonacute admissions calculated for nonacute days.

HOSPITAL VARIABLES:

(All variables represent hospital characteristics measured in fiscal year 1980.)

TOTLBEDS:	A continuous variable equal to hospital bed size.
EXPBED:	A continuous variable equal to the hospital's 1980 expenditures divided by TOTLBEDS.
TEACH:	Coded 1 if the patient is affiliated with a medical school or has a residency program and coded 0 if the hospital is a nonteaching hospital.
OCCRATE:	A continuous variable equal to the hospital's occupancy rate.
PROFIT:	Coded 1 for proprietary hospitals and 0 for nonprofit and nonfederal government hospitals.
PROFIT*OCC:	A term representing the interaction of PROFIT and OCCRATE.

AREA CHARACTERISTICS:

HBEDPERK:	A continuous variable equal to the number of short term general hospital beds per capita in 1981 in the FIPS county area.
MDSPERK:	A continuous variable equal to the number of patient-care physicians per capita in 1979 in the FIPS county area.
NBEDPERK:	A continuous variable equal to the number of nursing home beds per capita in 1978 in the FIPS county area.
PREV_CHG:	The average prevailing charge in 1978 for 12 major inpatient procedures in the FIPS county area.
POVERTY:	The percentage of families below the poverty level in 1970 in the FIPS county area.
POPOLD:	The percentage of the population over 74 years old in 1980 in the FIPS county area.

NBEDOLD: The ratio of nursing home beds (NPEDPERK) to persons over 74 years old (POPOLD) in the FIPS county area.

OLDPHY: The percentage of physicians over 55 years old in 1979 in the FIPS county area.

RURAL: Coded 1 of the hospital is located in a non-SMSA. Coded 0 otherwise.

N_EAST: Coded 1 if the hospital is located in the Northeastern region. Coded 0 otherwise.

SOUTH: Coded 1 if the hospital is located in the Southern region. Coded 0 otherwise.

WEST: Coded 1 if the hospital is located in the Western region. Coded 0 otherwise.

EXHIBIT V-10B

HOSPITAL-LEVEL INTERCORRELATIONS AMONG HOSPITAL AND AREA VARIABLES
(n=253)

	TOTLBEDS	EXPBED	TEACH	OCCRATE	PROFIT	PROFIT*OCC	HBEDPERK	MDSPERK	NBEDPERK	PREV_CHG	POVERTY	POPOLD	NBEDOLD	OLDPHY	RURAL	N_EAST	SOUTH
EXPBED	.48***																
TEACH	.67***	.43***															
OCCRATE	.51***	.47***	.33***														
PROFIT	-.20**	-.10	-.23***	-.23***													
PROFIT*OCC	-.18**	-.07	-.22**	-.14*	.97***												
HBEDPERK	.27***	-.01	.22**	.07	-.08	-.08											
MDSPERK	.48***	.34***	.39***	.24***	-.06	-.05	.59***										
NBEDPERK	-.13*	-.04	-.04	-.01	.00	.00	-.44***	-.22**									
PREV_CHG	.38***	.55***	.14**	.29***	.08	.07	-.07	.33***	-.10								
POVERTY	-.20**	-.32***	-.18**	-.29***	.08	.08	.22**	-.15*	-.25***	-.18**							
POPOLD	.05	.04	-.06	.13*	-.07	.07	.24***	.04	.08	.11	-.03						
NBEDOLD	-.12	-.02	-.07	.04	.02	.02	-.35***	-.20**	.98***	-.07	-.24**	.39***					
OLDPHY	-.21**	-.16*	-.30***	-.07	-.08	-.09	.10	-.34***	.13*	-.00	.01	.58***	.33***				
RURAL	-.44***	-.43***	-.40***	-.30***	-.02	-.02	.04	-.35***	.03	-.33***	.47***	.28***	.16**	.45***			
N_EAST	.01	.23**	.10	.22**	-.09	-.07	-.14*	.14*	.06	.27***	-.22**	.15*	.12	.06	-.11		
SOUTH	-.04	-.37***	-.11	-.09	.19**	.20**	.19**	-.14*	-.17**	-.27***	.58***	-.12	-.20**	-.12*	.15*	-.37***	
WEST	-.03	.24***	.05	-.07	.06	.04	-.18**	.07	.07	.34***	-.13*	.02	.06	.10	-.09	-.22***	-.33***

* = significant at the .05 level

** = significant at the .01 level

*** = significant at the .0001 level

direction, and, additionally, is negatively associated with location in the South (r = -.37) and positively with location in the West (r = .24).

Occupancy rate is highly correlated with not only bedsize and expenditures per bed, but teaching status (r = .33). Rural areas have fewer large hospitals (r = -.44), teaching hospitals (r = -.40), lower occupancy rates (r = -.30), fewer physicians per capita (r = -.35), lower prevailing charges (r = -.33) and have more people under the poverty level (r =.47), a higher percentage of the population over 75 years old (r = .28), and more older physicians (r = .45). The South has lower hospital expenditures per bed (r = -.37), a lower prevailing charge (r = -.27), and higher poverty levels (r = .58). There are more hospital beds per capita in areas with more physicians per capita (r = .59) and fewer in areas with more nursing home beds per capita (r = -.44).

Because of the intercorrelations among the variables, a forced-order regression analysis was performed in order to examine the contribution of certain variables, controlling for the effects of others. Variables were entered in the order shown above. Thus, the effect of resource availability (EXPBED) was considered only to the degree it contributed above and beyond size (TOTLBEDS). That for teaching status (TEACH) was of interest only to the degree that it contributed above and beyond size and resource intensity, regional differences (N_EAST, SOUTH, WEST) were of interest only to the degree they contributed above and beyond the effects of rurality (RURAL) and so forth.

Although NBEDOLD (or the ratio of nursing home beds to the elderly population in an area) can be considered a variable measuring alternate care availability, it was introduced after POPOLD. It was hypothesized that the higher morbidity of a large elderly population might be associated with a lower proportion of nonacute admissions. The opposite effect associated with a smaller supply of nursing home beds for this population might appear only after controlling for the first effect.

Exhibit V-11A presents a summary of the general linear regression analysis performed for the eighteen variables. Presented first is the sign of the estimated coefficient upon entry of the variable into the equation and second is the F value for the variable as it is entered into the equation. Variables were entered into the equation in the same order as listed in the exhibit, so this value indicates whether the variable makes a significant contribution to the variance above and beyond those variables previously entered. Listed third is the beta coefficient for each variable's contribution to the model as a whole, and the F value for the variable's contribution to the model. This value measures the effect of each variable above and beyond all other variables in the equation. Presented last is the zero-order correlation of each variable with the percentage of nonacute admissions.

As shown in column six of Exhibit V-11A, there are nine variables with a significant bivariate relationship with the percentage of nonacute admissions (as indicated by the zero-order correlation coefficient). These are hospital bedsize (TOTLBEDS), expenditures per bed (EXPBED), occupancy rate (OCCRATE), teaching status (TEACH), physicians per capita (MDSPERK), percentage of the population below the poverty level (POVERTY), rural versus urban location (RURAL), location in the Southern region (SOUTH), location in the Northeast (N_EAST) and location in the West (WEST). The direction of these effects is consistent with those found in the descriptive analyses and the patient-level logit model. Namely, the percentage of nonacute admissions is higher for nonteaching hospitals, hospitals with fewer beds, hospitals with lower expenditures per bed, and for hospitals located in rural areas, in areas with fewer physicians per capita, in areas with higher poverty levels, and in the South. The percentage of nonacute admissions is lower in the West and Northeast.

As shown in column three of Exhibit V-11A, only six variables make a significant contribution to the equation when entered successively. These are bedsize (TOTLBEDS) expenditures per bed (EXPBED), nursing home beds per capita (NBEDPERK), the percentage of older physicians (OLDPHY), and the three regional

EXHIBIT V-11A

HOSPITAL LEVEL ADMISSION ANALYSIS: HOSPITAL AND AREA CHARACTERISTICS GLM REGRESSION MODEL

VARIABLE (1)	SIGN OF COEFFICIENT UPON ENTRY (2)	F VALUE UPON SUCCESSIVE ENTRY INTO EQUATION (3)	COEFFICIENT FOR FINAL MODEL (4)	F VALUE FOR VARIANCE CONTRIBUTION TO MODEL (5)	ZERO ORDER CORRELATION OF VARIABLE WITH PERCENTAGE NONACUTE ADMISSIONS (6)
Intercept			0.116*		
TOTLBEDS	-	10.77***	-0.00005*	5.41*	-.19*
EXPBED	-	8.80**	-0.000	2.30	-.24**
TEACH	0	0.00	-0.002	0.49	-.15*
OCCRATE	0	0.19	-0.000	1.72	-.16*
PROFIT	0	1.80	-0.090	2.72	-.03
PROFIT*OCC	0	0.93	0.000	1.03	-.03
MDSPERK	0	0.45	-0.000	0.00	-.15*
HBEDPERK	0	0.23	-0.003	1.84	-.02
NBEDPERK	-	3.91*	0.017	2.20	-.09
PREV_CHG	0	1.72	0.000002*	10.50**	-.10
POVERTY	0	0.85	-0.000164*	3.90	.17*
POPOLD	0	0.00	-0.000	0.00	-.02
NBEDOLD	0	0.67	-0.090	3.84	-.08
OLDPHY	-	4.86*	-0.0009*	4.13*	-.01
RURAL	0	1.35	0.018	2.49	.19*
N_EAST	-	8.45**	-0.033*	8.71**	-.17*
SOUTH	+	16.15***	0.033*	9.35**	.27**
WEST	-	4.80*	-0.026*	4.80*	-.15*

F for Model = 3.66***
Model R-Square = .220

* = significant at .05 level
** = significant at .01 level
*** = significant at .001 level

dummy variables. Although expenditures per bed is negatively associated with nonacute admissions even when controlling for bedsize, none of the other hospital characteristics, including teaching status and occupancy rates, make a contribution above and beyond these factors. The effect of physicians per capita, in addition, is not significant after controlling for these variables, although an effect of nursing home beds per capita does emerge. The latter effect is in the predicted direction, i.e., fewer nursing home beds per capita are associated with more nonacute admissions. Another effect, not significant at a bivariate level, which emerges in the model is that for the percentage of older physicians, which is associated with fewer inappropriate admissions after controlling for other hospital and area variables. Even after introducing all other hospital and area variables, the effects of the three regional dummy variables are significant, and are among the strongest effects in the model. The direction of these effects is consistent with that observed at the bivariate level.

Coefficients for the model as a whole are presented in the fourth column of Exhibit V-11A and the F test for the significance of the effect's contribution to the variance is presented in column five. The model is significant (F = 3.35, p< .0001) and accounts for 21% of the variance. Six variables contribute significantly to the model. These are bedsize (TOTLBEDS), prevailing charges for 12 major surgical procedures (PREV_CHG), the percentage of older physicians (OLDPHY), and all three regional dummy variables (N_EAST, SOUTH, and WEST). In the context of all variables in the model, the effect of nursing home beds per capita is no longer significant. Exexpenditures per bed and bedsize entered last. This was done to assess the effects that would be obtained under the assumption that nonacute utilization is a phenomenon arising from global practice style differences at the regional level and, secondarily, between urban and rural settings that drive what occurs at a hospital level. In addition, the global effects due to hospital size and resource intensity were considered of interest only to the degree that they contributed above and beyond more specific admissions, after controlling for all other factors in the model.

Because the regional variables had an effect beyond those of the more specific hospital and area characteristics with which they were correlated, another forced order regression was performed with the same eighteen variables, but with the regional variables and urban/rural location entered first. These were followed by the remaining area characteristics, and then by the more specific hospital financial and organizational variables, with expenditures per bed and bedsize entered last. This was done to assess the effects that would be obtained under the assumption that nonacute utilization is a phenomenon arising from global practice style differences at the regional level and, secondarily, between urban and rural settings that drive what occurs at a hospital level. In addition, the global effects due to hospital size and resource intensity were considered of interest only to the degree that they contributed above and beyond more specific organizational and financial differences such as ownership, occupancy rate, and teaching status.

Exhibit V-11B presents this alternate specification, showing the sign and significance of each effect as it is entered in the equation. As shown in Exhibit V-11B, both the effects for the Northeast and the South are significant, but that for the West is not. The effect for rural location is significant even after controlling for regional differences. As with the bivariate relationship, the direction of this effect is positive, indicating that rural hospitals tend to have more nonacute admissions than would be expected even from their more frequent location in the South. A significant effect of occupancy rate is obtained, even after controlling for the associated area level variables such as region and urban/rural location. Bedsize is also significant, as is expected, given that it is the last variable to be entered in the equation and was previously found to make a significant contribution to the model as a whole.

3. Summary

The descriptive analyses, and the logit patient-level model both yield similar conclusions. The patient and stay factors most associated with nonacute admissions are (1) the four diagnostic categories: digestive problems, mental disorders, musculoskeletal problems, and symptoms and ill-defined conditions; (2) placement in a critical care unit, (3) admission through an emergency room

EXHIBIT V-11B

HOSPITAL-LEVEL REGRESSION ANALYSIS: HOSPITAL AND AREA CHARACTERISTICS
GLM REGRESSION MODEL
ALTERNATE SPECIFICATION

VARIABLE	F VALUE OF VARIANCE CONTRIBUTION UPON SUCCESSIVE ENTRY INTO EQUATION	SIGN OF COEFFICIENT UPON ENTRY
N_EAST	9.54**	-
SOUTH	14.65**	+
WEST	3.19	0
RURAL	5.68**	+
OLDPHY	0.61	0
POVERTY	1.27	0
POPOLD	0.00	0
NBEDPERK	1.21	0
NBEDOLD	0.98	0
MDSPERK	2.72	0
HBEDPERK	3.44	0
PREV_CHG	2.04	0
PROFIT	3.40	0
OCCRATE	6.94**	-
PROFIT*OCC	1.17	0
TEACH	1.14	0
EXPBED	2.56	0
TOTLBEDS	5.41*	-

* = significant at .05 level

** = significant at .01 level

*** = significant at .001 level

(for other than elective surgery), (4) patients of younger age, and (5) the percentage of the stay spent in workup. Medicare pay source, however, emerges as a significant effect in the logit analysis, a fact which is consistent with the significant effect of patient age. In addition, in the descriptive analyses stays of three to five days had the highest rate of nonacute admissions, but the effect was not significant in the logit analysis due to differences in the way the variable was coded.

In the patient-level logit model, patient and stay characteristics had the largest effects in the model. Patient and stay variables alone produced a rank correlation of .48 between the predicted probability and the actual value. The rank correlation for hospital characteristics alone was .16, and that for area characteristics alone was .22. The model based on the combination of patient, hospital, and area characteristics has a rank correlation of .55. The hospital-level model which examines only hospital and area variables, although significant, accounts for only 22% of the variance. Thus, hospital and area effects are relatively small in both the hospital-level and patient-level model.

Although there are a number of hospital and area variables which are significant at a bivariate level, only a few load in the patient or hospital models due to the substantial collinearity among them. This collinearity makes it difficult to assign causal responsibility to various hospital and area factors. Only the effects of region, bedsize, older physicians and prevailing charges are significant when controlling for all other variables in the model. Although rural location is significant after controlling for region, it is not significant in the context of all eighteen variables. Similarly, nursing home beds per capita contributes after controlling for some hospital area variables, but does not have an effect when rural location and regional variables are entered in the equation. Occupancy rate also is significant after controlling for region and other area variables, but not after controlling for hospital size and expenditures per bed. The effect of expenditures per bed, in addition, appears to substantially overlap the regional effects, being significantly lower in the South and higher in the West and Northeast. The effects of older physicians and the prevailing charge rate, on the other hand, appear only after controlling for all or a substantial number of other variables.

C. FACTORS ASSOCIATED WITH NONACUTE DAYS

1. Descriptive Analyses

As with the analysis performed for admissions, the proportion of nonacute days was examined for a variety of patient, hospital, and area characteristics. The results are presented below. Unless otherwise noted the standard errors presented in the exhibits were calculated using the generalization function described in Chapter III. The function for obtaining the relvariance of any estimate of days is as follows:

$$V^2_{x'} = .0007454 + \frac{192,143}{x'}$$

(The reader may refer to Section B above for a discussion of how to use the formula in calculating tests of significance.)

a. Patient and Stay Characteristics Associated with the Rate of Nonacute Days

Exhibit V-12 presents the percentage of nonacute days for selected patient and patient stay characteristics, including expected source of payment, primary placement, patient age, patient disposition, the major ICDA class of the principal diagnosis, whether or not major surgery was performed during the stay, length of stay, phase of the stay, and day of the week. The results are discussed below.

i. Expected source of payment

As with the admission analysis, the only pay categories with a sufficient sample size to permit analysis were for Medicare, Medicaid, Blue Cross and other commercial insurance. As with that analysis, no differences of note were found among patients hospitalized under Medicare, Blue Cross, or other commercial insurance. For Medicaid patients, however, the percentage of nonacute days at 3.9% was substantially lower than that for Medicare (8.6%), Blue Cross (9.9%)

EXHIBIT V- 12

PERCENTAGE OF NONACUTE DAYS BY PATIENT AND STAY FACTORS

(Excluding All Days in Nonacute Admissions)

	r 7.8%		(N_N/N_T)	s.e.	
EXPECTED SOURCE OF PAYMENT					
Medicare	8.6%	*	(9,362/109,188)	0.93e	
Medicaid	3.9%		(657/16,679)	0.94e	
Other Government	6.5%	*	(205/3,146)	6.10	
Worker's Compensation	4.1%		(147/3,589)	4.59	
Blue Cross	9.9%		(3,739/37,645)	2.14	1.11e
Commercial Insurance	7.3%		(3,493/48,099)	1.64	
HMO	6.0%		(76/1,264)	9.27	
Self Pay Only	12.0%		(895/7,439)	5.23	
Other	12.4%		(479/3,855)	7.36	
Not Available	5.7%		(550/9,702)	3.25	
TOTAL			(19,603/240,606)		
PRIMARY PLACEMENT					
Medical	9.5%	*	(11,338/119,370)	1.18	
Surgical	7.5%		(6,531/87,206)	1.24	
Psych or Alcohol	7.0%		(679/9,709)	3.59	
ICU/CCU, Burn or Trauma	4.3%		(892/20,758)	1.95	
Other	0.3%		(8/2,237)	1.75	
Not Available	8.2%		(105/1,284)	10.60	
TOTAL			(19,553/240,664)		
PATIENT AGE					
16-29	6.4%	ns	(2,156/33,659)	1.85	
30-44	5.8%		(1,880/32,468)	1.80	
45-64	10.1%		(6,860/67,889)	1.60	
65-74	7.0%		(3,359/47,728)	1.62	
75+	9.1%		(5,276/57,762)	1.66	
Not Available	6.4%		(71/1,107)	10.21	
TOTAL			(19,602/240,613)		
PATIENT DISPOSITION					
Transferred to SNF	13.6%	*	(2,819/20,684)	3.30e	
Transferred to acute facility	10.4%		(547/5,251)	5.84	00.62e
Other subacute/Home Care	6.8%		(1,122/16,534)	2.71	
Terminal Cancer	15.2%		(699/4,584)	7.36	
Other Terminal	0.5%		(42/8,032)	1.11	
Left AMA	6.6%		(95/1,450)	9.01	
Routine	7.8%		(14,052/180,158)	0.87	
Not Available	5.7%		(224/3,919)	5.14	
TOTAL			(19,600/240,612)		

* = significant at the .05 level

EXHIBIT V-12 (continued)

	r 7.8%		(N_N/N_T)	s.e.
MAJOR ICDA CLASS OF PRINCIPAL DIAGNOSIS				
Musculoskeletal	14.6%		(2,073/14,253)	4.09
Symptoms	13.3%		(1,226/9,187)	4.92
Trauma	12.2%		(2,778/22,744)	3.01
Digestive	9.5%		(2,573/27,209)	2.46
Mental	8.9%		(1,120/12,652)	3.50
Neoplasms	8.6%	*	(1,837/21,397)	2.65
Pregnancy	6.9%		(696/10,115)	3.49
Respiratory System	6.9%		(1,440/20,841)	2.44
Genito-Urinary	6.4%		(1,178/18,317)	2.51
Circulatory	5.9%		(3,032/51,154)	1.45
Other	5.3%		(1,180/22,409)	2.07
Nervous System	4.5%		(469/10,334)	2.84
TOTAL			(19,602/240,612)	
MAJOR PROCEDURE				
Performed	7.2%	ns	(7,824/109,431)	1.08
Not Performed	9.0%		(11,777/131,150)	1.09
TOTAL			(19,601/240,581)	
LENGTH OF STAY				
1	3.2%		(99/3,135)	4.33
2	2.6%		(224/8,797)	2.33
3	5.6%		(650/11,613)	2.96
4	7.1%	*	(958/13,422)	3.08
5	8.0%		(1,149/14,441)	3.12
6	9.6%		(1,285/13,393)	3.53
7	10.8%		(1,449/13,382)	3.72
8-30	8.8%		(10,668/121,786)	1.12
31+	7.7%		(3,118/40,637)	1.83
TOTAL			(19,600/240,606)	
PHASE OF STAY				
Workup/Pretreatment	10.5%	*	(3,118/29,832)	2.46
Surgical Treatment Phase	0.7%		(267/40,029)	0.56
Medical Treatment Phase	6.6%	*	(7,888/120,339)	0.99
End of Stay	16.5%		(8,328/50,400)	2.29
TOTAL			(19,601/240,509)	
PHASE OF STAY: MEDICAL STAYS ONLY				
Workup/Pretreatment	13.1%		(1,627/12,420)	4.20
Surgical Treatment Phase	2.5%		(55/2,208)	4.60
Medical Treatment Phase	5.6%		(4,944/88,018)	1.08
End of Stay	18.1%		(5,151/28,529)	3.16
SUBTOTAL: Medical			(11,777/131,175)	
PHASE OF STAY: (continued)				

* = significant at the .05 level

EXHIBIT V-12 (continued)

	r 7.8%		(N_N/N_T)	s.e.
PHASE OF STAY: SURGICAL STAYS ONLY				
Workup/Pretreatment	8.6%		(1,491/17,415)	2.94
Surgical Treatment Phase	0.6%		(212/37,822)	0.53
Medical Treatment Phase	9.1%		(2,944/32,317)	2.22
End of Stay	14.5%		(3,177/21,875)	3.30
SUBTOTAL: Surgical			(7,824/109,429)	
GRAND TOTAL PHASE OF STAY			(19,601/240,604)	
DAY OF WEEK DISCHARGED (Day Before Discharge Only)				
Sunday	16.7%		(470/2,812)	10.07
Monday	18.7%		(699/3,729)	8.86
Tuesday	14.6%		(604/4,143)	7.60
Wednesday	16.0%	ns	(795/4,982)	7.19
Thursday	12.2%		(594/4,852)	6.52
Friday	10.5%		(580/5,552)	5.69
Saturday	13.7%		(629/4,611)	
DAY OF WEEK ADMITTED: SURGICAL STAYS (First Day Of Stay Only)				
Sunday	4.6%		(111/2,410)	5.92
Monday	6.5%		(206/3,181)	6.04
Tuesday	6.2%		(177/2,871)	6.22
Wednesday	3.1%	ns	(71/2,316)	4.97
Thursday	0.9%		(19/2,020)	2.98
Friday	2.2%		(24/1,062)	6.32
Saturday	4.1%		(60/1,456)	7.22
DAY OF WEEK ADMITTED: MEDICAL STAY (First Day of Stay Only)				
Sunday	3.4%		(59/1,742)	6.01
Monday	6.0%		(163/2,722)	6.30
Tuesday	4.9%	ns	(146/2,963)	5.51
Wednesday	6.4%		(149/2,325)	7.04
Thursday	3.4%		(92/2,698)	4.84
Friday	3.2%		(87/2,701)	4.71
Saturday	7.3%		(122/1,680)	8.78

r = ratio estimate of % nonacute admissions

N_N = estimated nonacute admissions

N_T = total estimated admissions

s.e. = standard error of r (estimated using the generalization function unless otherwise denoted by e, in which case standard error was calculated directly using the variances and covariances generated by the balanced repeated replication process).

ns = not significant

* = significant at the .05 level

NOTE: Brackets above denote that significance tests were conducted to compare the two ratios indicated by the bracket endpoints.

and Commercial Insurance (7.3%). The difference between Medicaid patients and these other three classes of patients combined was statistically significant.

ii. Primary placement

As shown in Exhibit V-12, there are no significant differences between patients placed in surgical units and those in medical units. The percentage of nonacute days for patients in critical care units (ICU, CCU, burn, shock, or trauma units), however, is significantly lower than that for patients in a general medical service. This may be due partially, however, to the fact that days spent in the special unit are automatically defined as appropriate by the criterion set used in evaluating days. Further analysis should focus on days of care once the patient is removed from the critical care unit, controlling for day in the stay. The present sample size does not, however, permit analysis at this level of detail.

No other significant differences relating to placement were found. Patients in alcoholism/chemical dependency units appeared to have a similar percentage of nonacute days to those in a surgical service. Sampling error for these patients was large, however, and did not permit analysis.

iii. Patient age

Unlike the findings with respect to admissions, there is no significant relationship between patient age and the percentage of nonacute days. In particular, there is no tendency for the elderly (both those 65 to 74 years old and those 75 years or older) to have either a higher or a lower percentage of nonacute days than younger patients.

iv. Patient disposition

If poor discharge planning or the unavailability of alternate care facilities contributes to the problem of unnecessary hospital utilization, this should be reflected in a higher proportion of nonacute days for those discharged to subacute care facilities than for those discharged routinely. Although sample size for these types of patients is small, the data do appear to support this hypothesis at least to a degree. As shown in Exhibit V-12, only 7.8% of the days for patients discharged routinely are nonacute, compared to 13.6% of the days for patients discharged to a skilled nursing facility. This difference is significant. When only days at the end of the stay are considered, these figures increase to 16% and 26% respectively. The latter difference, however, does not achieve significance. No similar tendency to have a higher rate of nonacute days is exhibited for patients discharged to other subacute care facilities or to home care.

v. Major ICDA class of principal diagnosis

As found for admissions, there is wide variation by body system with respect to the proportion of nonacute days. Among the body systems with the highest rates of nonacute days were the four identified previously as having high rates of nonacute admissions, namely, the musculoskeletal system (with 15% of the days being nonacute), symptoms and ill-defined conditions (13%), the digestive system (10%), and mental disorders (9%). In addition, the category of traumas, with 12% nonacute days, was among the top five classes in terms of nonacute days. It was suggested by the physician consultant that the nonacute days for traumas represent primarily nonpenetrating traumas.

vi. Performance or nonperformance of major surgery

Although there is a tendency for the percentage of nonacute days to be lower for surgical patients (7.2%) than medical patients (9.0%), this difference is not significant.

vii. Length of stay

Whereas patients who are hospitalized inappropriately tend to have short to moderate lengths of stay (3 to 5 days), the proportion of nonacute days for patients appropriately hospitalized appears to increase as the length of stay increases. Only around 3% of the days for one and two day stays are nonacute, compared to 10% of the days for six day stays and 11% of the days for seven day stays. Although the percentage of nonacute days falls for stays of eight days or more, the absolute number of nonacute days in these longer stays continues to increase. For example, 10.8% nonacute days in a stay of seven days represents an average of eight-tenths of a day, whereas 7.7% nonacute days in a stay of 31 days represents an average of 2.4 days.

Exhibit V-13, which presents the percentage of nonacute days in a stay by length of stay, explores this relationship further. As can be seen the number of stays having at least one nonacute day increases steadily as the length of the stay increases. Ninety-six percent of one and two day stays have no nonacute days, compared with only 70.4% for seven day stays. Thus, 30% of seven day stays contain at least one unnecessary day. As length of stay increases, the likelihood of having more than one inappropriate day also increases. Only 4% of the three day stays have two or more nonacute days, whereas 20% of the seven day stays do.

Exhibit V-14 presents the distribution of nonacute days by the length of stay and the day in the stay. It can be seen that, although nonacute days can occur on any day of the stay, the great majority tend to occur at the end of the stay. For three day stays, 47% of the nonacute days occur on the last day of the stay, with the other 53% evenly distributed over the first and second days. For four and five day stays, 43% and 49% of the nonacute days respectively occur on the last day. For six day stays, 39% of the nonacute days occur on the last day and 68% occur on the last two days. These numbers are 31% and 52% respectively for seven day stays. A similar trend is apparent for stays longer than 7 days, although these have been presented in collapsed form in Exhibit V-14.

EXHIBIT V-13

PERCENTAGE DISTRIBUTION OF PATIENTS BY LENGTH OF STAY AND NUMBER OF NONACUTE DAYS IN STAY (ACUTE ADMISSIONS ONLY)

LENGTH OF STAY (in days)	PERCENT OF NONACUTE DAYS IN STAY (figures are ratio estimates calculated on N)								N (in 1,000s)
	0	1	2	3	4	5	6	7	
	%	%	%	%	%	%	%	%	
1	96.4	3.6	-	-	-	-	-	-	3,063
2	96.4	2.1	1.5	-	-	-	-	-	4,275
3	91.7	4.4	0.9	3.0	-	-	-	-	4,049
4	83.4	10.2	2.8	1.6	2.1	-	-	-	3,430
5	74.9	14.5	6.6	1.0	1.6	1.5	-	-	2,747
6	76.6	8.0	9.0	2.1	3.5	0.7	0.1	-	2,192
7	70.4	9.6	7.7	6.1	1.5	1.3	0.2	3.3	1,921
8-14	82.9	-[1]	-	-	-	-	-	-	6,639
15-21	81.0	-	-	-	-	-	-	-	2,094
22-28	81.6	-	-	-	-	-	-	-	739
29+	87.9	-	-	-	-	-	-	-	961
TOTAL									32,112

[1] Sampling weights do not allow for exact estimates of number of nonacute days for stays longer than seven days.

N = estimated days

EXHIBIT V-14

PERCENTAGE DISTRIBUTION OF NONACUTE DAYS BY LENGTH OF STAY AND DAY IN STAY

LENGTH OF STAY (in days)	DAY IN STAY (figures are ratio estimates calculated on row N)										N
	1	2	3	4	5	6	7	8-13	14-30	30+	
	%	%	%	%	%	%	%	%	%	%	
1	100.0										98,722
2	47.8	52.2									223,897
3	26.7	26.2	46.5								650,048
4	18.2	15.8	21.1	43.4							958,477
5	10.8	9.0	9.7	21.2	48.6						1,149,193
6	6.1	3.4	8.4	14.4	29.3	38.5					1,285,241
7	11.1	6.8	8.4	8.9	12.7	21.5	30.6				1,449,206
8-13	8.6	5.0	4.3	4.8	7.0	9.9	14.3	45.9			4,520,201
14-30	2.9	2.7	2.7	2.8	3.0	3.8	6.3	26.2	49.7		6,148,260
30+	0.0	0.0	0.0	2.9	1.6	1.6	1.6	4.8	44.3	43.9	3,118,348
											19,601,593

N = estimated admissions

Beginning with five day stays, there is a noticeable tendency for the second to the last as well as the last day of the stay to have a higher proportion of nonacute days. It is not clear from the data presently available whether this "end of the stay" phenomenon affects an increasing number of days at the end of the stay as length of stay increases, or whether unnecessary utilization at the end of the stay tends to be confined to the last two days.

The remainder of the nonacute days appear to be evenly distributed over the stay, and do not occur disproportionately in the beginning of the stay or the middle of the stay.

viii. Phase of stay

Exhibit V-12 next presents the proportion of nonacute days by phase of the stay. The categories shown in Exhibit V-12 are based upon the six treatment phases discussed previously: namely, "workup", "pretreatment", "surgical treatment", "medical treatment", "transitional", and "treatment completed". In performing the analysis, "workup" and "pretreatment" were collapsed. The transitional phase (which has very few cases) was collapsed with the medical treatment phase. Rather than displaying results for the "treatment completed" phase, for which days are almost by definition nonacute, the proportion of nonacute days is shown for the end of the stay, a period which was defined as one to five days at the end of the stay depending upon the length of the stay. One and two day stays were not considered as having an "end of stay" phase.

Not surprisingly, in light of the previously discussed findings, the highest proportion of nonacute days, (17%) occurs at the end of the stay. At 11%, the workup/pretreatment phase has the next highest proportion of nonacute days. When workup and pretreatment are analyzed separately, days in a workup phase were considered nonacute 16% of the time, versus 10% for pretreatment. The category "workup," therefore, has almost the frequency of nonacute days as the end of the stay.

Days occurring during the medical treatment phase were considered nonacute about 7% of the time whereas days spent in the surgical treatment phase were nonacute less than 1% of the time. The surgical treatment phase includes the day of the procedure and the following post-operative observation period. Thus, the provision of an acute level of service, combined with the patient's need for continued hospitalization following the procedure will almost automatically ensure that these days are appropriate.

Exhibit V-12 shows phase of stay further broken down for medical and surgical cases. The same relative distribution across phases is shown for both types of patients: the proportion of nonacute days is highest at the end of the stay, followed by the workup phase, and lowest in the surgical treatment phase. Although it appears that the percentage of nonacute days is higher for medical than surgical stays in both the workup and end of stay phase, these differences are not statistically significant. An interesting difference between medical and surgical cases appears, however, when the workup and pretreatment phases are analyzed separately. For medical cases, 13% of the days in a workup phase and 23% of the days in a pretreatment phase are nonacute. For surgical cases, 19% of the days in a workup phase are nonacute, but only 5% of the days in the pretreatment phase. Thus, the pretreatment phase for surgical cases appears to involve a different style of care than the workup phase; however, the same may not be true for medical cases.

ix. Day of the week

Preliminary analyses performed on the entire sample of days showed no significant differences in the proportion of nonacute days by day of the week. Further analyses were then undertaken to identify any effects of day of the week that might occur for the first and last days of the stay. In particular, we hoped to ascertain whether there was any tendency for patients kept over the weekend and discharged on Monday to exhibit a higher than average proportion of nonacute days on the last day of the stay. Similarly, the analysis sought to establish whether the first day of the stay was more often nonacute for surgical patients admitted on Friday than for patients admitted on other days.

Results of this analysis are presented in Exhibit V-12, which gives the proportion of nonacute days by day of the week for the first day of the stay and the day before discharge. The former is presented separately for medical and surgical stays. With respect to the day before discharge, Exhibit V-12 shows that the highest proportion of nonacute days (18.7%) does occur the day before a Monday discharge. This appears to support the supposition that Monday discharges may be more inefficient than discharges occurring on other days. With respect to the first day of the stay, on the other hand, the lowest proportion of nonacute days occurs for surgical patients admitted on Friday or Saturday. This runs counter to the hypothesis that Friday and Saturday admissions for surgery have a greater tendency to be inefficient. None of these differences is significant, however, due to the small sample sizes on which the figures are based.

To examine this issue further, Exhibit V-15 presents the distribution of hospital stays by the day of the week of admission and discharge. The day of week of admission is shown separately for admissions for an elective procedure versus all other admissions. The distribution of stays by day of discharge is broken down separately for stays with a major procedure versus those without, although no particular effect of this variable on the day of discharge was hypothesized.

As shown in Exhibit V-15, discharges are not evenly distributed by day of week as would be expected if they were completely determined by the patient's condition. The highest proportion of discharges occurs on Friday (20%) and the lowest on Sunday (9%) indicating that the day of discharge is somewhat dictated by administrative convenience. It is unclear, however, whether discharges for the approximately 6% additional patients who would normally be discharged on Sunday are being delayed or whether they are actually being hastened. The high rate of Friday discharges would tend to support the latter supposition. In addition, as can be seen in Exhibit V-15, there is no tendency for Monday to have a high discharge rate, suggesting that there is no noticeable tendency to hold patients for a Monday discharge.

EXHIBIT V-15

PERCENTAGE DISTRIBUTION OF HOSPITAL STAYS BY DAY OF THE WEEK FOR ADMISSION AND DISCHARGE

	ALL STAYS	STAYS FOR ELECTIVE PROCEDURE	STAYS NOT FOR ELECTIVE PROCEDURE
(figures are ratio estimates calculated on the column N)			
DAYS OF THE WEEK OF ADMISSION			
Sunday	12.3%	19.2%	11.1%
Monday	19.1%	22.3%	18.3%
Tuesday	18.1%	19.3%	17.5%
Wednesday	15.0%	17.8%	14.2%
Thursday	14.9%	24.5%	14.5%
Friday	12.0%	4.6%	13.8%
Saturday	8.1%	2.3%	10.7%
N	34,001	7,072	26,928
DAY OF THE WEEK OF DISCHARGE		STAYS WITH MAJOR PROCEDURE	STAYS WITHOUT MAJOR PROCEDURE
Sunday	8.6%	8.9%	9.4%
Monday	11.1%	10.5%	13.6%
Tuesday	15.2%	12.7%	14.2%
Wednesday	16.0%	17.8%	14.8%
Thursday	14.4%	16.3%	15.4%
Friday	19.6%	17.7%	18.5%
Saturday	15.2%	16.2%	14.2%
N	34,001	18,732	15,269

N - estimated admissions (in 1,000s)

The evidence with respect to premature admissions for surgery is also ambiguous. As shown in Exhibit V-15, the admission rate is highest on Monday and Tuesday (19% and 18% respectively) and declines steadily over the course of the week, with the smallest percentage of admissions occurring on Saturday (8%). These trends, however, vary markedly for elective admissions versus other admissions. For admissions for elective procedure, there is a fairly constant rate of admissions (ranging from 18% to 25%) from Sunday through Thursday. This declines to 5% on Friday and 2% on Saturday. Other types of admissions show some decrement on Friday, Saturday, and Sunday but do not have nearly the variation by day of week as elective admissions. It, therefore, appears that for elective admissions there is a deliberate attempt to schedule admissions so as to minimize weekend hospital days prior to the procedure. Nevertheless, 6.9% of elective admissions (representing 1.4% of all admissions) did occur on Friday or Saturday. These may potentially represent a source of inefficiency. Without knowing the day on which the procedure was performed, however, it is difficult to draw any conclusions.

Exhibit V-16 further shows the distribution of stays by the day of week of admission and the number of days from the day of the admission to the day of the procedure. This table includes not only admissions for elective procedures but all admissions in which a procedure was performed. An examination of the marginal column distribution shows that the distribution of stays by the admission day of the week is similar to that shown in Exhibit V-15 for elective procedures only. That is, there is a high rate of Sunday admissions and a substantial drop-off in Friday and Saturday admissions. As can be seen, in the majority of cases (50%) the procedure was performed the day after admission, and in 21% it was performed the same day. The procedure was performed two days later for 8% of the cases and three days later in 5% of the cases. In 15% of the cases the primary procedure was not performed until four or more days into the stay.

In Exhibit V-16, a dotted box is used to indicate an admission for which the procedure was performed on a weekend. For every category, the percentage of patients having a procedure performed on the weekend is substantially lower than those having a procedure on the weekday. For patients having the procedure

EXHIBIT V-16

PERCENTAGE DISTRIBUTION OF ADMISSION BY THE DAY OF THE WEEK AND NUMBER OF DAYS PRIOR TO THE PRIMARY PROCEDURE

DAY OF WEEK OF ADMISSION	PROCEDURE PERFORMED SAME DAY AS ADMISSION	PROCEDURE PERFORMED ONE DAY AFTER ADMISSION	PROCEDURE PERFORMED TWO DAYS AFTER ADMISSION	PROCEDURE PERFORMED THREE DAYS AFTER ADMISSION	PROCEDURE PERFORMED MORE THAN THREE DAYS AFTER ADMISSION	ROW N (Column %)
	(figures are ratio estimates calculated on column N)					
Sunday	8.6%	21.0%	18.8%	27.1%	10.9%	2,484,418 (17.0%)
Monday	14.9%	22.7%	23.8%	17.2%	20.8%	3,008,304 (20.6%)
Tuesday	15.5%	19.9%	17.8%	22.8%	16.5%	2,690,913 (18.4%)
Wednesday	16.8%	17.4%	24.2%	6.0%	12.9%	2,425,309 (16.6%)
Thursday	13.6%	15.7%	2.7%	1.9%	18.6%	2,029,909 (13.9%)
Friday	18.3%	2.3%	3.2%	12.7%	11.2%	1,117,333 (7.7%)
Saturday	12.2%	0.9%	9.5%	12.4%	9.2%	855,527 (5.9%)
COLUMN N (Row %)	3,117,251 (21.3%)	7,360,456 (50.4%)	1,221,524 (8.4%)	740,658 (5.1%)	2,171,824 (14.9%)	14,611,713 (100.0%)

performed one day after admission, only 2.3% were admitted on Friday for Saturday surgery and 1.9% on Saturday for Sunday surgery. For patients having a procedure performed two days following admission, only 2.7% were admitted on Thursday for Saturday surgery and only 3.2% on Friday for Sunday surgery. Similarly, for patients having a procedure performed three days after admission these numbers are 6.0% and 1.9% respectively. These percentages--ranging from about 1% to about 6% with an average of 3.9%--might represent the irreducible minimum of surgeries whose timing is not discretionary but must be performed as urgent or emergent interventions. It is interesting to note that when the procedure is performed the same day as the admission (column one in Exhibit V-16) there is much less of a decrement for Saturday and Sunday procedures, suggesting that this category is apt to include a large number of patients who require immediate surgical intervention.

In Exhibit V-16, a solid black line indicates Friday admissions for Monday surgery and Saturday admissions for Tuesday surgery. These account for 12.7% and 12.4% respectively of all patients for whom the procedure was performed three days after admission. Assuming approximately 4% each day were patients whose procedures could not be postponed, this means that about 9% of the Friday admissions for Monday surgery and 8% of the Saturday admissions for Tuesday surgery potentially could have been scheduled in a more efficient manner. It is also possible, of course, that most or all of these patients required three pre-operative days for workup or other pre-operative preparation.

Small sample sizes make it difficult to ascertain which of these possibilities is correct. However, it is clear that in magnitude, this problem accounts for a relatively small proportion of all nonacute hospital days. In all, there were 740,658 patients who had their primary procedure performed three days following admission. Assuming that 8.8% of the Friday admissions for Monday surgery and 8.5% of the Saturday admissions for Tuesday surgery were premature, there are 128,134 patients in all who may have had up to two unnecessary pre-operative days. This amounts to a maximum of 256,268 days or a total of only one tenth of a percent of all days of care, and only 1% of the 19,599,006 nonacute days of care occurring in otherwise justified admissions.

Although Friday admissions for Monday surgery do not appear to represent a particular problem in and of themselves, inefficient scheduling of surgery as a whole may still contribute significantly to unnecessary hospital utilization. For patients having major surgery during the stay, 8.6% of the 17,414,803 days spent in a workup or pretreatment phase were judged to be unnecessary. This amounts to 1,491,000 days, and represents almost 1% of all days of care, as well as 8% of nonacute days of care.

b. Hospital Characteristics and the Rate of Nonacute Days

Exhibit V-17 presents the proportion of nonacute days by a number of hospital characteristics including bedsize, ownership, teaching status, occupancy rate, and hospital expenditures per admission. Unlike the admission-level analysis, which found significant effects of bedsize, teaching status, occupancy rate, and expenditures per bed, the only characteristic showing a significant relationship with the proportion of nonacute days is occupancy rate. As with the admission analysis, hospitals with the highest occupancy rates (over 89%) have the lowest percentage of nonacute days and, conversely, the highest percentage of nonacute days occur for hospitals with the lowest (less than 60%) occupancy rates. The difference between these hospitals in the proportion of nonacute days is more than a factor of two: 13% versus 6%. The trend is not completely consistent, however, since hospitals with occupancy rates in the range of 80% to 89% actually have a higher proportion of nonacute days than those with occupancy rates in the 60% to 79% range (9% versus 7%). The latter difference is not statistically significant.

c. Area Characteristics and the Rate of Nonacute Admissions

Exhibit V-18 presents the percentage of nonacute days by selected area characteristics, including region, SMSA size, nursing home beds per capita, hospital beds per capita, physicians per capita, and the percentage of the population below the poverty level. The only variable among these exhibiting a significant relationship with the proportion of nonacute days is region. Unlike the admission analysis, which showed high rates of nonacute admissions in the South and low rates in the West and Northeast, the day level analysis reveals

EXHIBIT V-17

PERCENTAGE OF NONACUTE DAYS BY HOSPITAL CHARACTERISTICS (Excluding All Days in Nonacute Admissions)

	r 7.8%		(N_N/N_T)	s.e.
BEDSIZE				
6-99	9.4%		(2,915/31,029)	2.30
100-299	7.7%	ns	(6,047/78,145)	1.32
300-499	10.1%		(6,828/67,785)	1.60
500+	6.0%		(3,812/63,633)	1.30
OWNERSHIP				
Government, Nonfederal	9.8%		(4,728/48,392)	1.87
Nonprofit	7.7%	ns	(13,315/172,309)	0.89
Proprietary	7.8%		(1,559/19,899)	2.64
TEACHING STATUS				
Medical School Affiliated	7.6%		(6,457/85,240)	1.26
Residency Only	12.2%	ns	(2,028/16,632)	3.52
Nonteaching Hospital	8.0%		(11,116/138,729)	1.01
OCCUPANCY RATE				
0-59%	13.0%		(3,004/23,144)	2.19e
60-79%	6.8%		(6,689/99,045)	1.10
80-89%	9.0%	*	(8,450/93,590)	1.30
90%+	5.9%		(1,459/24,810)	1.37e
HOSPITAL EXPENDITURES PER ADMISSION				
0-\$54,807	9.0%		(4,639/51,816)	1.74
\$54,807-69,696	8.1%		(4,761/58,896)	1.56
\$69,697-85,231	7.2%	ns	(4,510/62,296)	1.44
\$85,232+	8.4%		(5,692/67,584)	1.48

r = ratio estimate of % nonacute admissions

N_N = estimated nonacute admissions

N_T = total estimated admissions

s.e. = standard error of r (estimated using the generalization function unless otherwise denoted by e, in which case standard error was calculated directly using the variances and covariances generated by the balanced repeated replication process).

ns = not significant

* = significant at the .05 level

NOTE: Brackets above denote that significance tests were conducted to compare the two ratios indicated by the bracket endpoints.

EXHIBIT V-18

PERCENTAGE OF NONACUTE DAYS BY AREA CHARACTERISTICS
(Excluding All Days in Nonacute Admissions)

	r 7.8%		(N_N/N_T)	s.e.
REGION				
Northeast	9.4%	*	(5,494/58,587)	1.27e
North Central	8.1%		(5,784/71,889)	1.41
South	8.3%		(6,405/76,957)	1.38
West	5.8%		(1,919/33,158)	1.60e
SMSA SIZE				
Non-SMSA	8.6%	ns	(4,401/51,158)	1.72
SMSA of <100K to 1 million	7.2%		(61,501/85,017)	2.13
SMSA > 1 million	8.7%		(9,049/104,426)	1.21
NURSING HOME BEDS PER CAPITA (x1,000)				
0-4.0	8.8%	ns	(5,884/66,641)	1.52
4.1-7.0	7.4%		(8,108/109,543)	1.10
7.0+	8.7%		(5,610/64,417)	1.54
HOSPITAL BEDS PER CAPITA (x1,000)				
0.3-4.0	9.5%	ns	(6,394/67,022)	1.57
4.0-6.0	9.2%		(8,730/94,541)	1.31
6.0+	5.7%		(6,210/78,017)	1.34
PHYSICIANS PER CAPITA (x1,000)				
0.08-1.2	7.9%	ns	(6,145/77,945)	1.34
1.21-2.0	8.6%		(7,247/84,629)	1.33
2.01+	8.0%		(4,478/79,029)	1.14
PERCENTAGE OF POPULATION BELOW THE POVERTY LEVEL				
0-7.5	9.1%	ns	(6,941/76,274)	1.44
7.5-10.0	10.0%		(5,468/54,953)	1.77
10.0-15.0	6.2%		(4,347/70,203)	1.26
15.0+	7.3%		(2,845/39,165)	1.82

r = ratio estimate of % nonacute admissions

N_N = estimated nonacute admissions

N_T = total estimated admissions

s.e. = standard error of r (estimated using the generalization function unless otherwise denoted by e, in which case standard error was calculated directly using the variances and covariances generated by the balanced repeated replication process).

ns = not significant

* = significant at the .05 level

NOTE: Brackets above denote that significance tests were conducted to compare the two ratios indicated by the bracket endpoints.

the highest proportion of nonacute days to be in the Northeast (9.4%) and the lowest to be in the West (5.8%). This difference is significant.

Just as the pattern of nonacute admissions by region followed the observed pattern of admissions per capita by region, the pattern of nonacute days mirrors the observed regional variation in average length of stay: namely, both nonacute days and average length of stay are highest in the Northeast, lowest in the West, and intermediate in the South and North Central regions.

d. Reviewer-supplied Reasons for Nonacute Utilization

Exhibit V-19 presents the breakdown of reviewer-supplied reasons for the occurrence of a nonacute day. These are shown both overall, and broken down by the phase of the stay, which includes "Workup/Pretreatment," "Medical Treatment", and "End of the Stay." "Surgical Treatment" was omitted since it represents the day of surgery and the immediate post-operative observation period and, therefore, very rarely results in a nonacute day.

Reasons are presented first for days judged nonacute because there is no need for continued hospitalization and second for days judged nonacute because the level of service is not appropriate. The great majority of nonacute days (over three-fourths) met neither a level of service criterion nor a need for continued hospitalization criterion. These days are classified in Exhibit V-19 under the heading "No need for continued hospitalization." Patients who met a level of service criterion but did not meet a need for continued hospitalization criterion are also classified under this heading. The category "Continued Hospitalization Needed; Services Not Acute Level" includes only patients who still require hospitalization, but are not provided services in a timely manner. This includes all patients with major surgery during the pre-operative portion of their stay. These pre-operative days are assigned to the "Workup/Pretreatment" phase. The reader may refer to Exhibit III-5 in Chapter III for a list of the reasons used by reviewers.

EXHIBIT V-19

DISTRIBUTION OF REASONS FOR NONACUTE DAYS BY THE PHASE OF THE STAY

REASON FOR NONACUTE DAY	PHASE OF THE STAY			
	WORKUP/ PRETREATMENT	MEDICAL TREATMENT	END OF THE STAY	TOTAL STAY
	(figures are ratio estimates calculated on column Ns)			
No Need for Continued Hospitalization:				
No Further Care Required:				
Terminal patient	0.2%	0.2%	0.1%	0.1%
Pt, MD insistence	7.0%	12.8%	8.8%	10.6%
Delayed D/C orders	1.9%	3.9%	9.8%	5.9%
Pending Patient Ed	0.0%	2.9%	0.4%	1.3%
No Home Care	1.5%	4.2%	1.1%	2.4%
Other/Not Known	34.8%	24.5%	30.2%	28.6%
Subtotal	44.4%	48.5%	50.4%	48.9%
Subacute Care Required:				
Terminal patient	1.7%	2.7%	4.2%	3.1%
Condition Improving	3.7%	3.1%	13.8%	7.7%
No alternate care	5.5%	14.1%	10.9%	11.2%
Delayed D/C orders	1.8%	3.8%	5.8%	4.3%
Pt, MD insistence	6.1%	9.4%	2.2%	5.7%
Other/Not Known	4.6%	10.0%	5.0%	6.9%
Subtotal	23.4%	43.1%	41.9%	38.9%
Continued Hospitalization Needed; Services Not Acute Level:				
Tests unavailable	13.8%	2.7%	1.9%	4.1%
Change in patient cond.	3.3%	1.1%	1.4%	1.6%
Patient noncompliance	2.2%	0.4%	1.0%	0.9%
Resource availability	1.8%	1.8%	0.2%	1.3%
Sequencing/Delayed orders	7.5%	0.4%	0.8%	1.7%
Other/Not Known	3.7%	2.1%	2.5%	2.7%
Subtotal	32.3%	8.5%	7.8%	12.3%
N	3,124,117	7,888,246	8,321,887	19,334,250[1]
ROW %	15.9%	40.2%	42.5%	100.0%

N = total estimated nonacute days

[1] Includes all nonacute days occurring in acute admissions except for an estimated 267,000 nonacute days occurring during the active surgical phase.

Exhibit V-19 shows that, in all, 16% of the unnecessary days occur in a workup or pretreatment phase, 40% in a medical treatment phase, and 43% at the end of the stay. An examination of the last column shows that only 12% of the unnecessary days occur because of the untimely delivery of services (i.e., when continued hospitalization is needed but the level of services is nonacute). For 88% of the days, the day was judged to be nonacute because the patient did not require hospitalization at all. On 49% of these, the reviewer felt that the patient did not require institutional care of any kind. On another 39%, the reviewer felt that the patient didn't require hospital care but did require some form of subacute care.

The fact that on 39% of the days reviewers felt the patient required a lower level of care suggests that a substantial proportion of nonacute hospital utilization may occur because of the unavailability of alternate care. Within this category of patients, however, reviewers felt that 4.3% of the problems occurred because of delayed discharge orders, and 5.7% because of patient, family or physician insistence. Another 7.7% represented patients who remained hospitalized because their condition was improving and recovery was anticipated within three days. For 6.9% of these days reviewers could not identify a reason that the patient remained hospitalized. From this category, then, there remain only 14.3% of nonacute days (11.2% where no alternate care is available and 3.1% for terminal patients) which the reviewers could definitely attribute to the lack of alternate care.

Two categories among patients not requiring either institutional or subacute care might also be considered "placement problems." These are the 0.1% terminal patients retained in the hospital for humanitarian reasons and the 2.4% for whom no home care is available. When added to the previous number, this yields an estimate that 16.8% of all nonacute days are due to lack of alternate care. To be conservative one might also add to this the 6.9% of days on which the reviewer could not identify the reason a patient requiring subacute care was retained in the hospital. This yields an estimate of 23.7% of nonacute days occurring because of the unavailability of alternate care, a figure which might be considered an upper bound.

Of the 49% of patients not requiring institutional care of any kind, reviewers attributed responsibility to patient or physician insistence in 10.6% of the cases and were unable to assign responsibility in 28.6% of the cases. A Chi-square test was performed to compare nonacute days in this latter category (i.e., those for whom no reason could be deduced) with those attributed to patient or physician insistence. These classes were compared with respect to rural/urban location (defined as non SMSAs versus SMSAs), patient age (over 75 years versus under 75 years), and pay source (Medicaid, self-pay, no charge, and other government versus all others). No significant differences between the two groups were found.

The relatively low percentage of nonacute days attributable to the inefficient provision of services was unexpected. It had been hypothesized that this would be the major factor contributing to nonacute days occurring in the "workup/pretreatment" phase. Additionally, it was hypothesized that this factor would be relatively important in explaining nonacute utilization in the medical treatment phase, although relatively unimportant for the end of the stay. Service-related reasons are indeed a larger factor in explaining nonacute days for the workup/pretreatment phase than they are for any other phase: 32% of the days in a workup phase are inappropriate for service-related reasons compared to around 9% for the medical treatment phase and 8% at the end of the stay. Still, with the inefficient provision of services accounting for only 32.3% of the unnecessary days in the workup/pretreatment phase, the over-riding problem is that the patient did not require the acute care setting at all for performance of the workup. In 24% of the cases, reviewers felt that a subacute setting was required and, in 44% of the cases, that no institutional care was needed at all and the workup could be provided on an outpatient basis.

A very similar pattern is found for the medical treatment phase. In most cases, a day is inappropriate not because of delays in the provision of services but because the patient does not require the acute care setting at all. With respect to the end of the stay, not surprisingly, most nonacute days occur because the patient does not require further hospitalization, rather than because of the inefficient provision of services.

2. Multivariate Analyses

As with the admission level-analyses, two sets of multivariate analyses were performed to predict nonacute days. The first was at the patient-level and the second at the hospital-level. These are discussed below.

a. Patient-level Model

The logit regression procedure was again used to build a model for predicting nonacute days at the patient level. To avoid the problem of correlations among observations arising from the similarity of days in the same stay, only one day was selected from each stay. This day was selected by random assignment with all days having equal probability of being selected. If the length of stay was only one day, that day was included in the sample with certainty. In the analysis, days were weighted by the original sampling and nonresponse weights, rather than by the ratio estimation weights.

As with the day analysis, three separate stepwise regressions were first performed: one for patient characteristics only, one for hospital characteristics only, and one for area characteristics only. The variables used in this analysis were presented previously in Exhibit V-7.

Exhibit V-20A presents the stepwise regression analysis for the patient variables. As shown in Exhibit V-20A, eight patient variables entered significantly into the equation through the stepwise regression. These are the same variables that were significant prior to entry of any of the variables into the equation (as shown in the right hand column). The significant effects are largely in accordance with the descriptive findings. A day occurring at the end of the stay (ENDSTAY) is substantially more likely to be nonacute. The same is true of a day occurring in the workup phase (WORKUP). These are among the largest effects in the equation. Patients discharged to a skilled nursing facility (SNFDISP), and those hospitalized for mental disorders (MENTAL) or for symptoms and ill-defined conditions (SYMPTOMS) are also more likely to have nonacute days. Patients having shorter lengths of stay (SHORTLOS) and those placed in ICUs and other critical care units (SPECPLAC) are less likely to have

EXHIBIT V-20A

PATIENT-LEVEL DAY ANALYSIS: PATIENT CHARACTERISTICS STEPWISE LOGIT REGRESSION MODEL

All Variables Prior to Entry into Model:			Variables Entered Through Stepwise Regression:		
VARIABLE	BETA	CHI-SQUARE	VARIABLE	BETA	CHI-SQUARE
PATAGE	+	13.53*	Intercept	-3.38	836.17*
CAIDPAY	0	0.36	ENDSTAY	1.25	91.61*
CAREPAY	+	3.67	WORKUP	1.29	68.67*
SPECPLAC	-	10.86*	PROC	0.68	28.51*
URGADMIT	0	0.99	SPECPLAC	-1.46	23.04*
SYMPTOMS	+	14.54*	SHORTLOS	-0.79	17.30*
DIGPROB	0	0.57	MENTAL	0.30	8.68*
MUSKEL	0	0.05	SNFDISP	0.61	6.57*
MENTAL	+	2.11	SYMPTOMS	0.46	4.64*
PROC	+	52.56*			
WORKUP	+	81.33*			
ENDSTAY	+	111.69*			
TERMPAT	0	0.16			
SNFDISP	+	14.06*			
SHORTLOS	-	37.48*			
LONGLOS	0	0.30			

Model Chi-Square = 294.25*

Rank Correlation between Predicted Probability and Response = .50

* = significant at the .05 level.

nonacute days. One variable which has a very strong effect in the model, but not in the descriptive analysis, is the performance of a major procedure during the stay (PROC) which is negatively associated with the probability of a nonacute day. Since the appropriateness of the level of service and the need for continued hospitalization are given almost automatically on the pre-operative preparation day, the day of the procedure, and at least two days following the procedure, this is not a surprising finding, and we expected the relationship to be significant to the descriptive analysis. Although the effect was not significant at the descriptive level, the observed differences were in the appropriate direction.

The patient-level stepwise regression model is significant (Chi-square = 294.25, $p< .0001$) and the rank correlation of the predicted and actual values is .50. The model was cross-validated using a second random sample of days. All significant effects were replicated with the exception of that for MENTAL which, although significant, was not as strong an effect as that attained in the original model. In addition, MUSKEL, which is not significant in the original model, was significant in the cross-validation model.

Exhibit V-20B presents the stepwise regression equation calculated for hospital characteristics. Although only occupancy rate is significantly related to the proportion of nonacute days at the descriptive level, three variables contribute significantly to the equation. These are TOTLBEDS, OCCRATE, and PROFIT. The direction of the effects is such that profit hospitals and hospitals with more beds have less probability of a nonacute day, while hospitals with higher occupancy rates have a higher probability of a nonacute day. The latter contradicts the relationship found in the descriptive analysis, although it should be noted that in the descriptive analysis this trend was not entirely consistent.

In a second random sample of days, the occupancy rate effect was replicated. Those for bedsize and profit/nonprofit status were not. Expenditures per bed however, were negatively associated with the probability of a nonacute day. This variable is highly correlated with bedsize and may have substituted for bedsize in the equation. The hospital characteristics stepwise

EXHIBIT V-20B

PATIENT-LEVEL DAY ANALYSIS: HOSPITAL CHARACTERISTICS STEPWISE LOGIT REGRESSION MODEL

All Variables Prior to Entry into Model:			Variable Entered Through Stepwise Regression:		
VARIABLE	BETA	CHI-SQUARE	VARIABLE	BETA	CHI-SQUARE
TOTLBEDS	-	4.36*	Intercept	-3.096	74.25*
PROFIT	-	3.27	TOTLBEDS	-0.001	10.31*
OCCRATE	+	2.02	OCCRATE	0.001	4.50*
PROFIT*OCC	-	2.09	PROFIT	-0.326	4.22*
EXPBED	0	0.18			

Model Chi-Square = 15.67*

Rank Correlation between Predicted Probability and Response = .12

* = significant at the .05 level

logit model is significant (Chi-square = 15.7, $p < .001$) with a rank correlation of .12.

Exhibit V-20C presents the stepwise regression model for area characteristics. Only two variables contributed significantly to the equation: namely, the two dummy variables for the western (WEST) and southern (SOUTH) regions. Both were associated with a lower probability of nonacute days. These effects were replicated in a second random sample of days. The model is again significant (Chi-square = 22.2, $p < .001$) but the rank correlation is only .07.

Exhibit V-21 presents a logit regression model containing all the patient, hospital, and area variables previously examined. The same effects noted above are significant with the exception of those for occupancy rate and profit/nonprofit status. Both of these drop from the equation. Bedsize, however, continues to make a fairly strong contribution. Regional effects due to the South and West dummy variables are also highly significant and in the expected direction.

The overall model is significant (Chi-square = 358.82, $p < .0001$) and the rank correlation between the predicted probability of a nonacute day and the actual value is .55.

b. Hospital-level Model

Exhibit V-22 presents the hospital level model for the day level analysis. This model was constructed in the same manner as that for the admission analysis described previously, and the same weighting technique was used. Shown in Exhibit V-22 is (1) the value of the estimated coefficient (Beta), (2) the F value for the contribution of the effect to the equation, (3) the F value for the effect as it is entered into the equation in the order given in the exhibit, and (4) the zero-order correlation between the variable and the percentage of nonacute days.

EXHIBIT V-20C

PATIENT-LEVEL DAY ANALYSIS: AREA CHARACTERISTICS
STEPWISE LOGIT REGRESSION MODEL

All Variables Prior to Entry into Model:			Variables Entered Through Stepwise Regression:		
VARIABLE	BETA	CHI-SQUARE	VARIABLE	BETA	CHI-SQUARE
N_EAST	+	3.09	Intercept	-2.40	1069.08*
SOUTH	-	3.62	WEST	-0.74	17.23*
WEST	-	11.63*	SOUTH	-0.38	10.31*
RURAL	0	0.79			
HBEDPERK	0	1.09			
NBEDPERK	+	2.98			
MDSPERK	0	0.80			
POPOLD	0	0.13			
POVERTY	0	1.78			

Model Chi-Square = 22.82*

Rank Correlation between Predicted Probability and Response = .07

* = significant at the .05 level

EXHIBIT V-21

PATIENT-LEVEL DAY ANALYSIS: PATIENT, HOSPITAL, AND AREA CHARACTERISTICS LOGIT REGRESSION MODEL

VARIABLE	BETA	CHI-SQUARE
Intercept	-2.77	21.63*
PATAGE	0.00	0.33
CAIDPAY	-0.03	0.02
CAREPAY	-0.25	1.91
SPECPLAC	-1.56	24.67*
URGADMIT	-0.31	2.84
SYMPTOMS	0.57	6.71*
DIGPROB	0.11	0.40
MUSKEL	0.13	0.32
MENTAL	1.02	11.03*
PROC	0.73	28.71*
WORKUP	1.52	81.03*
ENDSTAY	1.03	51.34*
SHORTLOS	-0.74	18.16*
LONGLOS	-0.30	3.76
TOTLBEDS	-0.001	11.57*
PROFIT	-0.32	0.16
OCCRATE	0.00	1.37
PROFIT*OCC	-0.00	0.00
EXPBED	-0.00	0.13
N_EAST	-0.24	1.72
SOUTH	-0.51	6.10*
WEST	-0.78	11.37*
RURAL	-0.28	1.98
HBEDPERK	-0.06	2.72
MDSPERK	0.13	2.85
NBEDPERK	0.02	0.94
POPOLD	-0.04	0.68
POVERTY	0.00	0.73

Model Chi-Square = 358.82*

Rank Correlation between Predicted Probability and Response = .55

* = significant at the .05 level.

EXHIBIT V-22

HOSPITAL-LEVEL DAY ANALYSIS: HOSPITAL AND AREA CHARACTERISTICS
GLM REGRESSION MODEL

VARIABLE	BETA	F VALUE FOR VARIANCE CONTRIBUTION TO EQUATION	F VALUE UPON SUCCESSIVE ENTRY OF VARIABLE INTO EQUATION	ZERO ORDER CORRELATION OF VARIABLE WITH PERCENTAGE NONACUTE DAYS
Intercept	.213 *			
TOTLBEDS	-.000	1.08	1.02	-.00
EXPBED	.000	0.01	0.00	.00
TEACH	-.002	0.19	0.32	.01
OCCRATE	-.000	0.77	0.06	.06
PROFIT	.023	0.06	2.05	-.08
PROFIT*OCC	-.000	0.50	0.81	-.09
HBEDPERK	-.006	2.73	2.86	-.11
MDSPERK	.005	0.57	0.47	-.05
NBEDPERK	-.025	2.43	1.19	.00
PREV_CHG	.000	0.50	0.06	.05
POVERTY	-.000	0.72	0.20	-.09
POPOLD	-.007	1.22	0.11	.01
NBEDOLD	.005	1.73	2.69	.03
OLDPHY	-.001	0.78	0.78	.04
RURAL	.018	0.92	0.91	.03
N_EAST	-.006	0.09	0.23	.04
SOUTH	.004	0.05	0.98	-.03
WEST	-.040	3.98*	3.98*	-.06

F for Model = 1.04
Model R-square = .07

* = significant at the .05 level

As can be seen in Exhibit V-22, although one variable (WEST) was significant in the model, the model itself failed to achieve significance. Although the effect for WEST is in the predicted direction, that variable was not significant at a bivariate level.

3. Summary

As in the admission analysis, the strongest effects in the day analysis are patient and stay characteristics. The logit patient-level model for patient characteristics produces a rank correlation between predicted and actual values of .50.

The patient variables significant in descriptive analysis are consistent with those found in the logit model analysis. Namely, days at the end of the stay, days in a workup phase, days for patients discharged to a skilled nursing facility and days for patients with longer lengths of stay are more likely to be nonacute.

Patients receiving surgery and those receiving critical care are less likely to have nonacute days (the former effect emerging strongly in the logit analysis but not in the descriptive analysis). For both these classes of patients, however, certain days of active treatment are almost by definition acute-level. Sample size did not allow us to evaluate whether there are any differences between surgical, critical care, and medical patients in the less active portion of the stay.

As found in the admission analysis, four diagnostic categories are associated with nonacute days: these are digestive problems, mental disorders, musculoskeletal problems, and symptoms and ill-defined conditions. In addition, traumas, which are not associated with nonacute admissions do have a higher proportion of nonacute days. In both the descriptive and logit analyses, the effects of diagnostic category were somewhat unstable.

Hospital and area characteristics have far less of an association with nonacute days than they do with nonacute admissions. The rank order correlation of the logit model probabilities with actual values is .12 for hospital characteristics and .07 for area characteristics. (The rank correlation of the complete logit model containing patient, hospital, and area characteristics is .55, compared with .50 for patient characteristics alone.) Similarly, the GLM regression performed at the hospital level did not achieve significance at all.

Only occupancy rate and region are associated with nonacute days at the descriptive level. In the logit model, the hospital effects are unstable. Occupancy rate is significant in the stepwise regression submodel for hospital characteristics but the effect is in the opposite direction from that in the descriptive analysis. Bedsize and expenditures per bed (which are highly correlated) variously emerge in different models. The regional effects, however, are stable and in the expected direction. In the hospital-level model, region makes a significant contribution, but the model itself is not significant.

VI. SUMMARY AND CONCLUSIONS

VI. SUMMARY AND CONCLUSIONS

Previous efforts to study the problem of unnecessary hospital utilization have relied on comparisons of utilization rates themselves across geographical boundaries or practice settings. Studies which have measured inappropriate hospital utilization directly have used unreliable methods and/or have been based on small and non-representative samples. Accordingly, there has been much controversy over the magnitude of unnecessary hospital utilization and its causes.

This is the first study which has applied a single, criterion-based methodology for measuring inappropriate hospital utilization in a nationally-representative probability sample. This methodology, termed the Standardized Medreview Instrument, was specifically developed for the project from a synthesis of existing utilization review techniques. It is capable of assessing both the necessity of the admission and of each day of care.

After extensive field testing, the instrument was applied in a sample of 253 hospitals nationwide to examine the records of patients discharged in 1981. Information on 5,732 admissions and 24,051 days of care was abstracted. The purpose of the study was to obtain reliable estimates of the level of unnecessary hospital utilization and to identify the factors associated with such utilization. The major findings of the study are summarized here.

A. THE MAGNITUDE OF NONACUTE HOSPITAL UTILIZATION

Using the Standardized Medreview Instrument, the study estimated that 5.7% of the admissions for patients discharged in 1981 were unnecessary, and 12% of the days of care. Of the latter, 4.2% were unnecessary because the admission itself was unnecessary and 7.8% were unnecessary days that occurred in otherwise justified admissions.

The Standardized Medreview Instrument is based on criteria established and in widespread use for PSRO review at the time the study was initiated. In the course of instrument development the criteria were enhanced to incorporate revisions to the original criteria set that had been made by over 50 PSROs. These revisions were reviewed and integrated by a physician panel. The criteria were further streamlined and clarified based upon the results of a fifteen-hospital field test, and the modifications again reviewed by the physician panel. In the survey, reviewers achieved more than acceptable levels of reliability in applying the criteria. In addition, the reviewer evaluations were extensively investigated following data collection and were found to have been made in accordance with the intention of the criteria set.

An independent evaluation of the Standardized Medreview Instrument and its properties relative to both other utilization review methods and established medical judgment is a high priority in the continued effort to understand the phenomenon of nonacute hospital utilization. However, based on the lengthy development and evaluation process this instrument, underwent, we feel (1) that the criteria of the Standardized Medreview Instrument are in accordance with accepted standards of medical practice and utilization control and (2) that the criteria were applied by reviewers in a manner consistent with their intent.

The findings of this study -- that 5.7% of admissions and 12% of hospital days are unnecessary -- indicate that there is not as alarming a degree of hospital misutilization as is often suspected. Although this study did not disclose blatant inefficiency in the use of hospital resources, the magnitude of unnecessary hospital use found here should not be viewed as trivial. One hundred eighteen billion dollars were spent on hospital care in 1981. If 12% of hospital days of care could be eliminated and/or redirected to lower cost settings, the potential cost savings are immense. Some inefficiencies in the delivery of health care are inevitable, of course, and cannot be controlled without incurring even greater economic and societal costs. Nevertheless, the findings of the study -- to be discussed below -- indicate that the majority of unnecessary care could be eliminated through changes in the practices of hospitals, physicians, and patients.

The 5.7% unnecessary admissions and 12.0% unnecessary days of care found in this study should not be taken to imply that there are no cost savings to be gained beyond this level. First, the study in no way touched upon the cost savings that might be possible through more efficient use of ancillary services in the hospital. Second, the necessity of elective surgeries was not assessed in this study. Nineteen percent of the admissions in the study were for major (i.e., Class 1) surgery which was defined as elective in that the patient did not otherwise meet a criterion signifying the need for immediate intervention. If even 5% of these are unnecessary, the overall number of inappropriate admissions is increased by almost one percent.

Finally, and more centrally, the criteria of the Standardized Medreview Instrument might be considered as defining the upper boundary of what McClure (1982) terms the "range of acceptable practice." Within this range, as McClure points out, there are physicians, such as those in many prepaid group practices who make "conservative use of inpatient physician and hospital services" while others make aggressive use of these services. The magnitude of unnecessary hospital use found in this study, thus, does not reflect the reductions possible from practicing a more "conservative" style of medicine or those to be gained from widespread implementation of newly developing alternatives to hospital-level care. The 5.7% inappropriate admission rate, for example, does not include the 1.7% of procedures that were performed in the hospital, but which could have been performed in an ambulatory surgery unit if one were available.

The percentage of hospital admissions and days which could be eliminated is thus bound to increase as society searches for less costly alternatives to acute-level care and as these innovative approaches to alternate care become standard practice.

B. FACTORS ASSOCIATED WITH NONACUTE HOSPITAL UTILIZATION

As noted previously, analysis was aimed at addressing issues which had been raised in previous research on the subject of nonacute or unnecessary hospital utilization. These were as follows:

- what kinds of patient stays are most susceptible to nonacute use
- how much nonacute utilization is associated with inefficient scheduling or poor discharge planning
- how much nonacute utilization is associated with differences in hospital structure, resource availability, delivery style, management, and financial incentives
- how much nonacute utilization is associated with reimbursement mechanisms (especially insurance coverage) which encourage overuse of hospital care
- how much nonacute utilization is due to the unavailability of subacute care
- how much nonacute utilization is associated with area factors which might create induced demand
- how much of the regional variation in utilization rates (i.e., admission rates and ALOS) can be attributed to nonacute utilization

The following summarizes the major findings of the admission-level and day level analyses with respect to these issues.

1. Major Findings of the Admission-level Analysis

Although there are a number of patient and patient-stay characteristics which are highly associated with nonacute admissions, these tend not so much to identify the population "at risk" for an inappropriate admission as to reflect the medical circumstances of the case which defined it as nonacute in the first place. Thus, admissions with characteristics indicative of more serious or life-threatening problems are less likely to be inappropriate. These include patients placed in a critical care unit upon admission, as well as patients admitted through the emergency room versus those admitted routinely (after excluding those admitted for elective surgery). Older patients are also less likely to be inappropriately admitted, most likely because of the more serious nature of their illnesses. Similarly, patients with long lengths of stay (reflecting more serious problems) tend to have inappropriate admissions less frequently.

Inappropriate admissions conversely tend to occur in the non-vital organ systems, with the smallest portion falling in the class of circulatory disorders. The four diagnostic classes with a noticeably large proportion of inappropriate admissions, on the other hand, are (1) digestive problems, (2) musculoskeletal problems, (3) symptoms and ill-defined conditions, and (4) mental disorders. In each of these categories, the medical problems characteristic of nonacute admissions tend to be symptoms or signs of undetermined etiology (e.g., migraine headache, back pain, gastroenteritis, syncope) which can be incapacitating and/or painful and can be indicative of serious or life-threatening conditions but which are not immediately life-threatening.

Ill-defined conditions of undetermined etiology representing no immediate threat to life undoubtedly account for the high proportion of admissions for workup which are inappropriate. This category had by far the largest rate of nonacute admissions, 12%, compared with 6.7% of patients admitted in a medical treatment phase. Although only 12% of the patients admitted in a workup phase are inappropriately admitted, they represent from 41% to 47% of all inappropriate admissions. (The exact number could not be determined because the first day of the stay was not abstracted for some stays.)

The use of an acute care setting to investigate medical problems which may be signs or symptoms of a condition potentially requiring acute-level treatment, but which do not in themselves warrant hospitalization, thus, appears to be a major factor in the overutilization of hospitals. Medical treatment which could be provided on an outpatient basis was the second largest factor, accounting for 33% to 38% of the unnecessary admissions.

There was no tendency for patients hospitalized inappropriately to be discharged quickly (i.e., within one or two days), although they did have shorter than average lengths of stay. The highest percentage of nonacute admissions occurred for stays of from three to five days. These stays, in addition, accounted for 45% of all inappropriate admissions.

In general, it is the physician who initiates hospital admissions, albeit with the acquiescense of or pressure from the patient and his family. One would, therefore, expect physicians and/or patients to be responsible for a significant portion of nonacute admissions. (Given the complexity of the physician-patient relationship, there was no attempt to differentiate physician from patient responsibility in the assignment of reasons for nonacute use.) Reviewers were able to attribute 41% of nonacute admissions definitely to physician/patient insistence. In another 41% of the cases, although the patient required no institutional care, reviewers were unable to assign a reason for the inappropriate admission. An examination of these patients showed no demographic differences between them and patients hospitalized because of physician or patient insistence. In particular, there was no tendency for the former to be older, to live more often in rural areas, or to be more disadvantaged (as indicated by pay source). It is interesting that in both the admission and day analysis, reviewers were most often unable to assign a reason when the patient did not need institutional care at all. For these patients, inability to assign a reason may mean that -- although there is no overt indication of patient or physician responsibility -- there is no good alternative explanation either.

There is, thus, a strong possibility that the majority of unnecessary admissions are controllable by physicians and their patients. There was only the most indirect evidence, however, of supplier-induced demand -- either of physician supply or the supply of hospital beds. Nonacute admissions were, in fact, significantly higher in areas with fewer physicians and hospital beds. McClure (1982) hypothesizes that the high utilization rates that occur in areas where physicians and hospitals are concentrated reflects -- not inappropriate utilization -- but utilization which lies within the range of acceptable practice. That is, the admissions which produce the higher admission rates are not medically inappropriate, but merely reflect more "elaborate" utilization philosophies which are nevertheless professionally defensible.

The professionally "nondefensible" portion of utilization identified by the Standardized Medreview Instrument did not, in fact, vary with per capita utilization rates at the county level at all (although it did vary with regional utilization rates). At the hospital level, admissions per capita in a FIPS county area had a -.02 correlation with the percentage of nonacute admissions. Similarly, average length of stay at the county level had a negligible correlation with the percentage of nonacute days.

In the Standardized Medreview Instrument, criteria are ordered from the most stringent to those where individual physicians could reasonably vary in terms of their decisions to hospitalize. It might be instructive in future analyses to examine variation in admissions by the stringency of the criterion to identify whether significant relationships with area-specific supply factors and utilization rates emerge. It is also possible, however, that the observed correlation of physician and hospital bed supply with utilization rates is accounted for by differences in elective surgery rates, in which case one would expect no effect of criterion stringency.

Nonacute admissions, then, varied inversely, rather than positively with physician and hospital supply variables, indicating that the factors promoting nonacute utilization are greater in low-supply areas than they are in high-supply areas. There was, however, one finding which suggested that when these influences are controlled, physician-induced demand may play a role in nonacute admissions. This was the relationship between nonacute admissions and the prevailing charge rate for twelve major procedures. This effect was significant only when controlling for all other variables in the model, and indicated that, all other things being equal, nonacute admissions are higher in areas where prevailing physician fees for twelve major surgical fees are higher.

Given that it is the physician who initiates the hospital admission, one would not necessarily expect hospital factors to be strongly related to the rate of unnecessary admissions. In this study, however, a variety of hospital characteristics were associated with higher rates of nonacute admissions. In particular, smaller hospitals, non-teaching hospitals, those with lower expenditures per bed, and those with lower occupancy rates had higher rates of nonacute admissions.

Larger hospitals, teaching hospitals, and those with higher expenditures per bed are generally known to treat the more severely ill. Thus, the observed differences could be accounted for by differences in case mix alone. However, even though small, nonteaching hospitals will naturally treat the less severely ill, this does not explain why such hospitals treat a higher proportion of the non-acutely ill.

The relationship between occupancy rate and unnecessary hospital admissions suggests that hospitals themselves may play a role in inducing inappropriate admissions. Occupancy rate, however, is correlated with a number of hospital and area factors and appears to have no effect above and beyond them. It is, therefore, difficult to draw any conclusions about the causal role of occupancy rates per se.

The hospital characteristics associated with nonacute admissions were also correlated with the area characteristics associated with nonacute admissions, indicating that the problems may not be hospital-induced but are endemic to an area. The characteristics associated with a higher rate of nonacute admissions were location in the South, in a rural area, in areas with fewer physicians per capita, higher poverty levels, older physicians (in some analyses), fewer nursing home beds per capita (in some analyses), and fewer hospital beds per capita (in some analyses). All these factors were correlated. That is, rural areas, which had few large hospitals, also had fewer physicians and hospital beds per capita, higher poverty levels, more older physicians, an older population, and were more often located in the South. This suggests that differences in nonacute utilization might be associated with a single regional or rural/urban dimension which reflects either differences in practice style between physicians in different geographic locations or differences in the availability of lower level facilities.

According to reviewer's judgments, up to 18% of the nonacute admissions were caused by the unavailability of subacute or home care. This figure includes the 3% of patients who required subacute care for whom reviewers were unable to deduce a reason for the cause of the unnecessary admission. Thus, a significant portion -- though not the majority -- of nonacute admissions are attributable to the lack of alternate sources of care. There was little evidence to suggest, however, that the higher nonacute admission rate found in rural areas was due to the lack of subacute care or that nonacute admissions were affected by the supply of nursing home beds at all. Nursing home beds per capita, although significant at a descriptive level, did not make an independent contribution to the model after controlling for all other variables. In addition, the supply of nursing home beds per capita was not significantly correlated with rural location.

Another possibility that might arise particularly in rural settings is that some patients who could be provided treatment or diagnostic workup on an outpatient basis are being hospitalized because of the unavailability of suitable outpatient facilities. Reviewer's notes indicated the unavailability of outpatient care to be a problem for only one admission in the data set. Nevertheless, the issue warrants further investigation. In performing the analysis, no good measure of outpatient treatment resources in the hospital area was available. Additional analysis will, therefore, require collecting this information for the sampled hospitals on a case-by-case basis.

The differences observed between small and/or rural hospitals in the proportion of nonacute admissions may also in part or as a whole be attributable to differences in patient expectations and physician practice styles. In the hospital-level model, the percentage of older physicians in an area was negatively associated with the rate of nonacute admissions, after controlling for hospital-specific factors, health resource supply factors, and other area factors. Very few physician characteristics were available for the analysis, and none were available at the patient-level. Although this relationship is quite possibly mediated by some unmeasured factor or is due to aggregation bias, it is one of the few available pieces of evidence to suggest that physician characteristics themselves may be related to differences in nonacute utilization.

The strongest evidence, however, that nonacute utilization was attributable to local differences in patterns and styles of care was the effect of region which emerged consistently in each analysis. This effect conformed to observed utilization patterns (i.e., high admission rates in the South and low in the West and Northeast) and was present even after controlling for a number of area and hospital factors which vary across regions. This would seem to argue, first, that observed regional variation in admission rates reflects regional tendencies toward overutilization and, second, that such tendencies are at least partly due to differences in practice style or philosophy, rather than to demographic and resource differences that vary across regions.

2. Major Findings of the Day-level Analysis

The patients most likely to incur nonacute days tended to fall in the same diagnostic categories that were associated with nonacute admissions: namely, musculoskeletal problems, digestive problems, mental disorders, and symptoms and ill-defined conditions. This was true even though the day-level analysis included only days for which the admission was appropriate in the first place. Thus, even patients appropriately admitted with symptoms, back problems and so forth were likely to incur a higher percentage of nonacute days. Additionally, in the day-level analysis patients with traumas (most likely nonpenetrating) were apt to have more nonacute days.

Not surprisingly, the types of stays most likely to have inappropriate days were longer stays, with the percentage of nonacute days increasing steadily as the length of stay increased from one to seven days. For stays over seven days, the proportion of nonacute days leveled off, but the absolute numbers of nonacute days in the stay continued to increase.

The end of the stay had the highest proportion of nonacute days (17%) and accounted for 43% of all nonacute days. The workup/pretreatment phase had the next highest proportion of nonacute days (11%), but accounted for only 16% of all nonacute days. Only 7% of the days occurring in a medical treatment phase were nonacute but these represented 40% of all nonacute days. The percentage of nonacute days in the active surgical treatment portion of a stay was negligible, accounting for only 1% of all nonacute days.

Surgical patients, defined as any patient having a Class 1 procedure performed during the stay, tended to have proportionately fewer nonacute days. It is not clear whether this was true of the end of the stay, but in the pretreatment and the active treatment phase, surgical cases had a definite advantage. Of all nonacute days, only 14% occurred prior to the day of the primary procedure for any patient having major surgery during his/her stay. Only 9% were for surgical patients in a treatment, pretreatment or workup phase (the remaining 4% being for patients in a medical treatment phase for whom surgery was subsequently scheduled). Of these, 5% were patients in a diagnostic workup phase, 3% were patients in the pretreatment phase, (i.e., those with the scheduled procedure pending), and 1% were patients in the active surgical phase.

Only 12% of all nonacute days were attributable to service delays (e.g., delays in ordering tests, scheduling surgery, receiving test results or consultations, etc.) which occurred when the patient otherwise required continued hospitalization, a category which included patients awaiting scheduled surgery. The highest percentage of days which were inappropriate for service-related reasons took place in the workup/pretreatment phase. Service-delays still accounted for only 32% of all the nonacute days occurring during this phase. Of these nearly half, representing 2% of all nonacute days, were attributable to service delays prior to surgery.

The great majority, 88%, of nonacute days were inappropriate not because of service and scheduling delays, but because the patient did not require acute level care. Seven percent of these were patients in a workup or pretreatment phase who required no institutional care of any kind. Presuming that (except for a negligible percentage of terminal patients and those with no home care) these are patients who could have had their pretreatment or diagnostic testing performed on an outpatient basis, this means that 7% of nonacute days could be eliminated through greater use of preadmission testing programs.

Some of these days might be attributed to inefficiencies which are inevitable when dealing with the disease process. Eight percent of nonacute days, for example, are for patients who require subacute care, but whose recovery is anticipated in such a short time that transfer to a subacute facility would not be worthwhile. Another 1.3% are delays pending patient education. Although these types of nonacute days (the latter category especially) might be avoided through more efficient scheduling of services and discharge, they may be difficult to eliminate within the margin of normal human error and efficiency.

Of the remaining 72% nonacute days, the patient required subacute care for nearly half. This suggests a substantial portion of nonacute days are due to the lack of availability of alternate care. Based upon reviewers' reasons, this did appear to be a major factor contributing to nonacute utilization. Counting the unavailability of home care, up to 24% of the nonacute days were attributed to the unavailability of alternate forms of care. (Included in this figure were all patients requiring subacute care for whom no reason could be deduced for the patient's continued hospitalization.) This observation is also consistent with the fact that patients discharged to a skilled nursing facility had more nonacute days than patients discharged routinely.

Of the 48% remaining days, however, nearly all appear to be potentially controllable. In about 15% of the cases (including 6% where the patient required subacute care) reviewers specifically attributed responsibility to patient or physician insistence. In another 10%, reviewers attributed the cause to delayed discharge orders. There were, finally, 23% of nonacute days in the medical treatment phase or at the end of the stay for which the patient did not require institutional care and reviewers were unable to assign a reason. The patients incurring these days were not significantly different demographically from those in which the patient/physician were assigned responsibility. That is, there was no evidence that these were overwhelmingly rural patients, elderly patients, or poor patients. Therefore, there is a strong indication that these unnecessary days are controllable and that they occurred through some combination of hospital, physician, or patient-caused inefficiency.

In sum, aside from 24% of nonacute days that reviewers attributed to the unavailablility of alternate care, and the 10% that were due to such factors as expected improvements in patient condition or patient education, around 66% of the nonacute days appeared to be potentially controllable. Seven percent were for diagnostic testing that should have been conducted on an outpatient basis prior to admission, and another 12% were due to delays in service, of which 2% were delays prior to scheduled surgery. Of the 66% potentially controllable nonacute days, 17% were attributed specifically to the physician and/or patient by reviewers, and 10% to delays in discharge orders. The latter are most likely some combination of physician and hospital responsibility.

Within this 66% of potentially controllable misutilization, there is also some question as to how close to perfect efficiency it is possible for the health care system to come. Some inefficiency in the planning and delivery of services is inevitable, particularly when dealing with a phenomenon as uncertain as the disease process. Nearly 1% of the delays in service, for example, were due to patient noncompliance or indecisiveness regarding the course of treatment, the latter of which may be considered a natural problem in human decision-making. Another 1.6% were due to changes in patient condition,[1] some of which may be difficult to anticipate in time to discharge the patient to home or subacute care. Similarly, other delays in discharging the patient may be difficult to control given the difficulty of both anticipating improvements in the patient's condition and coordinating the actions of the patient, his family, the physician, and the hospital staff.

---

[1] It should be noted that the latter do not represent days awaiting surgery when the patient's condition is in a changing or unstable status. These days met a level of service criterion (criterion number 28) and were considered appropriate. Days considered inappropriate due to a change in the patient's condition were those in which the patient improved to such a degree that he/she no longer required services and should have been discharged or days for a terminal patient, when the patient's condition deteriorated to such a point that acute intervention was no longer effective.

Many of the nonacute days, particularly those due to delays in service or delays in discharge are potentially the responsibility of the hospital. In addition, area characteristics, such as subacute care availability, may play a role in accounting for the 24% of nonacute days attributable to lack of alternate care. Unlike the admission analysis, however, there were few hospital and area characteristics which were correlated with nonacute days. There was some indication in the data that larger hospitals and/or those with higher expenditures per bed had fewer nonacute days. In addition, there was some indication that those with higher occupancy rates had fewer nonacute days. These relationships, however, were not particularly stable. Those relationships that were found at a descriptive level appeared to contradict those found in the logit analysis. In addition, none was significant in the hospital model.

This may reflect the fact that nonacute days are more under the control of individual physician practices than hospital and area factors. Alternatively, it is possible that hospital and area effects are operating in different directions for different phases of the stay. For example, nonacute days due to inadequate preadmission testing in the workup phase may be physician controlled and, thus, subject to the same influences that affect nonacute admissions. The nonacute days occurring in the pretreatment and treatment phases may be more influenced by hospital efficiency factors. The end of stay may be sensitive to area characteristics such as the number of nursing home beds per capita, and hospital efficiency factors may play a role as well. Future analyses might profitably attempt to fit separate models for different portions of the stay such as the workup phase, pretreatment phase, medical treatment phase, and end of the stay.

As in the admission analysis, the strongest and most consistent effects to emerge in the analysis of nonacute days were again regional factors. These effects were consistent with the observed regional pattern in average length of stay; namely, the rate of nonacute days was highest in the Northeast where average length of stay is highest and lowest in the West where average length of stay is lowest. This relationship, once again, suggests an effect of local practice style which is unassociated with any of the demographic and resource factors correlated with region.

3. Synthesis

The following summarizes the major implications of the study with respect to the eight research questions presented earlier.

a. What kinds of patient stays are most susceptible to nonacute utilization.

Forty-one to forty-seven percent of the unnecessary admissions were for diagnostic workup that should have been performed on an outpatient basis. Another 33% to 38% were for medical treatment that could have been provided at a lower level of care. Almost 60% of the nonacute admissions occurred in four body systems: digestive disorders, mental disorders, musculoskeletal problems, and symptoms and ill-defined conditions. Nonacute stays tended to be shorter than average: forty-five percent were from three to five days.

For stays up to seven days, the longer the stay, the more nonacute days it contained, both in absolute numbers and in terms of the proportion of nonacute days. For stays over seven days, the proportion of the stay that was nonacute leveled off, but the absolute number of nonacute days continued to increase. The last day of the stay or, for longer stays, the last few days had the highest probability of being nonacute, with the remainder of the nonacute days evenly distributed over the rest of the stay. Forty percent of nonacute days occurred at the end of the stay (defined as the last one to five days of the stay, depending on the length of stay); 40% occurred in a medical treatment phase; 1% in an active surgical treatment phase; and 20% occurred in a workup or pretreatment phase.

Together, the active surgical treatment phase and the surgical pretreatment phase (i.e., with scheduled surgery pending) had the lowest percentage of nonacute days and accounted for the least nonacute days. These two phases combined contributed only 4% of all nonacute days. Along with traumas (most likely nonpenetrating), the four conditions which were most apt to have nonacute admissions were also apt to have the most nonacute days; namely mental disorders, musculoskeletal problems, digestive problems and symptoms and ill-defined conditions. For nonacute days, however, there was no clear demarcation between body systems (in terms of high and low contributors) found for nonacute admissions.

b. How much nonacute utilization is associated with inefficient scheduling or poor discharge planning.

Very few nonacute days appeared to be caused by the inefficient scheduling of surgery. Only 3% of all nonacute days occurred when scheduled surgery was pending and, of these, only 2% were attributable to delays in service. Additionally the percentage of nonacute days that could be attributed to Friday admissions for Monday surgery was negligible. Only six-tenths of a percent of the patients in the sample were admitted on Friday and had major surgery performed on Monday. If all these patients (subtracting out a portion for which surgery was probably undelayable) had two unnecessary preoperative days, this would still amount to only 1% of all unnecessary days.

Only 12% of all nonacute days were due to delays in services during any phase of the stay. Five percent of these were in a workup or pretreatment phase and the remainder were distributed over the rest of the stay. Whether all of these could be eliminated is not clear. The very low percentage of nonacute days occurring during surgical pretreatment and active surgical treatment suggests that, once the treatment course is decided, hospitals function efficiently in providing services. A substantial number of service delays, therefore, may arise from lack of clarity as to the appropriate course of treatment.

Nonacute days arising from delays in discharge may, on the other hand, contribute substantially to the problem of nonacute utilization. There was, however, no evidence that patients held over the weekend for Monday discharge were more apt to have nonacute days (Due to small sample size, however, sampling error for the last day of the stay was large.) There was, in fact, a statistically higher than average percentage of discharges taking place on Friday and a statistically lower than average percentage of discharges taking place on Monday, suggesting that patients are more apt to be discharged on Friday than held over for Monday discharge.

Forty percent of all nonacute days, however, occurred at the end of the stay. Twenty-seven percent were not attributable to placement problems, changing patient condition, or any other extenuating circumstance. Reviewers specifically attributed 5% to physician or patient insistence and 10% to delayed discharge orders, but were unable to assign a reason to the other 12% of the nonacute days falling at the end of the stay, days for which the patient did not require subacute care. As with service delays, it may, nevertheless, be difficult to substantially increase the efficiency with which many patients are discharged.

In addition to delays in service or delays in discharge, another source of nonacute days which can be considered one of inefficient scheduling is that for diagnostic workup which could have been conducted on an outpatient basis. It appeared that up to 7% of nonacute days might be eliminated through a more efficient program of preadmission testing.

c. How much nonacute utilization is associated with differences in hospital structure, resource availability, delivery style, management practices, and organization.

The data offered no definitive answers to this question, but provided much evidence that was suggestive. In the admission analyses, in particular, a number of hospital characteristics were associated with nonacute admissions. Hospitals with low occupancy rates (less than 60%), for example, had a rate of nonacute admissions that was four times as high as hospitals with high (over 89%) occupancy rates. In addition, the trend was monotonic: the rate of nonacute admissions increased steadily as occupancy rate decreased. This strongly suggests that much nonacute utilization is due to demand that hospitals induce to compensate for low occupancy rates.

Large hospitals, teaching hospitals, and hospitals with higher expenditures per bed, however, also had lower rates of nonacute admissions. These attributes were not only highly intercorrelated, but had correlations with occupancy rate ranging from .33 (for teaching status) to .51 (for bedsize). Thus, case-mix differences, differences in practice style, and/or differences in efficiency might account for the observed differences between high and low occupancy rate hospitals. Conversely, differences in occupancy rates may account for the lower rate of nonacute admissions for large, teaching hospitals.

In an examination of these hospital characteristics, occupancy rate and teaching status contributed nothing above and beyond the effects of bedsize and expenditures per bed. Bedsize and expenditures per bed, moreover, were substantially correlated with area factors associated with nonacute admissions. In particular, large bedsize and higher expenditures per bed were negatively correlated with rural location (r= -.44 and -.43 respectively). Expenditures per bed was also highly correlated with region, being higher in the West (r = .23) and Northeast (r = .24) and lower in the South (r = -.37).

Region and rural location were not only associated with each other but with poverty level, nursing home beds per capita, physicians per capita, and hospital beds per capita. Few variables contributed after controlling for all others. Those that did, however, included both hospital characteristics (bedsize) and area characteristics (region). It appears, then, that hospital characteristics have some effect on nonacute utilization above and beyond area effects and, conversely, that the area influence on nonacute utilization is not solely mediated by the types of hospitals which predominate in an area. The complexity of the interrelationships among these variables suggests that a causal modeling approach which compares the explanatory power of variously specified models could profitably be employed to understand the mechanism through which various hospital and area characteristics interact in determining nonacute utilization.

Perhaps the most pertinent finding of the admission analysis was not the identification of which characteristics were related to nonacute admissions but the finding that at the hospital level, 22% of the variance in nonacute admissions was accounted for by hospital and area characteristics. It has been suggested that -- since the physician controls much of the decision-making in the hospital stay -- the factors contributing to nonacute use might be totally idiosyncratic and unpredictable, especially without knowing the characteristics of individual physicians and their patient clientele. Although 22% is not the majority of the variance, this finding nevertheless suggests that a nontrivial portion of nonacute utilization is due not to individual characteristics but to expectations, needs, demands, incentives or habits that operate at the macro-level to influence utilization decisions.

Unlike the admission analysis, the day analysis -- where one would expect hospital characteristics such as resource availability, financial incentives, and so forth, to have a large effect -- showed no consistent relationships of hospital-level variables with nonacute days. This may, on the one hand, indicate that hospitals are less responsible for nonacute utilization than physicians and their patients, whose characteristics were unmeasured in the analysis. On the other hand, hospital (and area) characteristics may play a strong role in inhibiting or contributing to nonacute utilization, but this role may vary depending upon the phase of the stay. It was suggested that more finely-tuned analysis is required in the future, which examines the effect of hospital and area characteristics for different phases of the stay.

Additionally, in the present study, missing data for some variables precluded analyzing some of the more precise measures of hospital resource availability and organizational structure (such as the type of utilization review department, the presence of a long-term care unit, organized rehabilitation unit, organized discharge planning unit and so forth). Such resources at a hospital level might significantly influence the efficiency with which care is delivered. An examination of these variables, however, would require obtaining detailed data directly from hospitals, rather than from secondary data sources. In addition, the effect of any single such organizational factor is bound to be small in relationship to all other factors affecting nonacute days. That is, the presence of a long-term care unit will promote the efficient discharge only of patients requiring subacute care, the presence of an organized rehabilitation department will affect only patients requiring rehabilitation services. In order to assess these effects, the sample from each hospital will have to be large enough to reliably represent the cases who would potentially be affected by a given organizational dimension. This would require a much larger sample from each hospital than obtained in the present study.

In this respect it is not surprising that in the day analysis the hospital-level model found no significant relationships between the percentage of nonacute days and any hospital or area characteristics. Given the small sample of records drawn from a hospital, the percentage of nonacute days and even admissions estimated for a given hospital will have a wide confidence interval. That is, the estimated percentage of nonacute days by hospital will be relatively imprecise. Any correlation coefficient which first aggregates the percentage of nonacute admissions or days by hospital and, then, examines the association of that variable with a particular hospital or area characteristic will be unstable. This implies that the descriptive analyses and the logit analyses, which aggregate the percentage of nonacute admissions or days for a particular hospital or area characteristic across all hospitals with that characteristic, will produce more stable results.

d. How much nonacute utilization is associated with reimbursement mechanisms (especially insurance coverage) which encourage overutilization of hospital care.

The attempt to examine the possible effects of different reimbursement mechanisms was hampered by the lack of necessary data. An examination of the rate of unnecessary utilization for self-paying patients was considered of interest, for example, under the assumption that such patients have incentives to curtail unnecessary utilization. The numbers of self-paying patients in the sample were too small, however, to draw any conclusions about their propensity or lack of propensity for unnecessary utilization. The utilization of patients covered by an HMO was similarly of interest and was similarly too small to analyze. In addition, the attempt to include a sample of HMO-owned hospitals in the study was unsuccessful. For patients covered by commercial insurance, a better knowledge of the details of the benefit structure, copayments and deductibles might have provided some insight into the effect of different incentives on the tendency to overutilize. However, this information was also not available.

The findings did indicate, however, that the Medicare and Medicaid programs are not subject to any untoward abuse with respect to unnecessary utilization. There was, in fact, no difference among the major payors (i.e., Medicare, Medicaid, Blue Cross, and other commercial insurance) in the proportion of nonacute admissions. There was, in addition, no difference among Medicare, Blue Cross, and other commercial insurance in the number on nonacute days. Medicaid patients, however, had a significantly lower proportion of nonacute days. This effect was not due to the younger age of Medicaid patients since age itself was uncorrelated with the proportion of nonacute days. This relationship was the only indication in the present study that the reimbursement practices of various payors might influence the overutilization of services.

e. How much nonacute utilization is due to the unavailability of subacute care.

According to reviewer judgments, up to 18% of the nonacute admissions and 24% of the nonacute days occurring in acute admissions may have been due to the unavailability of alternate care. The significantly higher proportion of nonacute days found for patients discharged to a skilled nursing facility also suggest that the unavailability of alternate care plays a nontrivial role in nonacute utilization.

Despite the magnitude of this problem, there was no strong association with the number of nursing home beds per capita found in the study, particularly for days of care. An effect of nursing home beds per capita was found for admissions but was correlated with other area variables and made no explanatory contribution above and beyond these variables. The apparent lack of effect may in part be due to the relatively small proportion of nonacute admissions or days attributable to alternate care problems coupled with the small sample of records obtained from each hospital. If the unavailability of alternate care is a factor in only one-fifth of the nonacute admissions or days, then instances of these may be unreliably indicated in any sample of 20 to 30 records. Correlations with area characteristics based on this sample of records will, thus, be unstable.

One aspect of the analysis which at first suggested that alternate care might be a major factor in nonacute utilization was the positive correlation found between nonacute admissions and both poverty level and rural location. These relationships suggest that the paucity of outpatient care as well as of subacute care could be a major contributor to the problem of nonacute admissions. There was no indication, however, that lack of outpatient facilities were a significant factor explaining either nonacute admissions or days. The effect of a rural location and high poverty levels, therefore, appeared to operate through some other mechanism.

f. How much nonacute utilization is associated with area factors that might create induced demand.

Rather than being positively associated with the number of physicians and hospital beds per capita, nonacute admissions were negatively associated with these factors. The number of physicians and hospital beds per capita were themselves positively associated with variables such as hospital bedsize, urban location, teaching status, and so forth, all of which had lower rates of nonacute admissions. If there was any tendency for the physician or bed supply to induce unnecessary admissions, then, it was masked by these countervailing trends. The only hint of an effect of supplier-induced demand in the analysis was the positive relationship between the prevailing charge rate and nonacute admissions which emerged after controlling for all the above influences.

It was also hypothesized that the higher utilization rates that occur in areas with higher physician supply represent utilization that is within the range of acceptable practice and, therefore, would not be defined as nonacute using the present methodology. Another possibility is that these higher utilization rates are comprised of elective surgeries whose appropriateness was not assessed in the present study.

g. How much of the regional variation in utilization rates (i.e., admission rates and average length of stay) can be attributed to nonacute utilization.

Although nonacute utilization was negatively associated with physician and hospital bed supply and had no asssociation with per capita utilization rates at a county level, they did appear to explain much of the regional differences in utilization rates. Aside from patient and stay characteristics, regional differences were the strongest and most consistent effects found in the study. The pattern of nonacute utilization found by region was consistent with observed variation in admission rates and average length of stay. Namely, the rate of nonacute admissions was highest in the South where admissions per capita are also highest and lowest in the Northeast where admissions per capita are also lowest. The rate of nonacute days was highest in the Northeast where average length of stay is highest and lowest in the West where average length of stay is also the lowest.

Regional effects were significant in explaining nonacute admissions even after controlling for a variety of hospital and area characteristics which were correlated with region. Regional effects were significant in explaining nonacute days even though there were no other hospital or area characteristics associated with nonacute days. This strongly suggests that regional differences in nonacute utilization represent global differences in practice style and are not mediated by other characteristics associated with region, such as resource availability, poverty, rurality, the types of hospitals predominating in the region, and so forth.

C. IMPLICATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

The major purpose of this study was to measure the magnitude of what we have termed nonacute hospital utilization, that is, the use of the hospital for other than the delivery of acute care. The second goal was to identify the factors associated with such utilization. Of specific interest was ascertaining how much of nonacute utilization can be eliminated through changes in the behavior of hospitals, physicians, and patients, rather than structural changes to the health delivery system.

In this national study, 5.7% of all admissions were found to be nonacute, a figure which does not include the possible contribution of unnecessary elective surgery or reductions in hospital admissions that might result from widespread adoption of more innovative practices in the delivery of care (such as the use of ambulatory surgery units for procedures which are now performed routinely on an inpatient basis, hospice care for the terminally-ill, and so forth).

Of these nonacute admissions:

- Up to 18% were patients who might have been inappropriately hospitalized because of the unavailability of alternate care (including patients with no home care and terminal patients hospitalized for humanitarian reasons)
- Another 6% were patients who needed subacute care, but were hospitalized at patient, family, or physician insistence.
- A total of 76% patients required no institutional care at all. Of these admissions, only 36% were definitely attributed by the reviewer to patient or physician insistence. For another 40% no reason could be deduced. However, these patients were not demographically different from those hospitalized at patient or physician insistence.

With respect to days of care, this study found that 12.0% of all days of care were nonacute. Of these 7.8% occurred in admissions that were justified. Within this group:

- Up to 24% were attributable to the unavailability of other forms of care.
- 9% represented discharges to subacute care which did not take place because of anticipated improvements in patient condition, or were for delays pending patient education.

The remaining 67% of nonacute days appeared to be potentially controllable:

- 12% were due to delays in service, tests, consultations and so forth, of which 2% took place while a patient with a scheduled procedure was awaiting surgery.
- 7% were for diagnostic testing that could have been performed on an outpatient basis
- 25% occurred at the end of the stay or in the medical treatment phase and were definitely attributed to physician/patient insistence or to delays in discharge order.
- 23% occurred at the end of the stay or in the medical treatment phase and were for patients who required no subacute care at all. Reviewers were unable to assign responsibilty for these nonacute days but these nevertheless appear to be under the control of the patient, physician or hospital in that these patients had no alternate care needs and were not demographically different from patients retained at physician or patient insistence.

Whether all these delays in service or in timely discharge can be eliminated or whether some represent inevitable inefficiencies inherent in the medical decision process is unclear.

Although this study found a lower rate of inappropriate utilization than has generally been accepted, the amount is still nontrivial in terms of costs to both federally- and privately-funded health care programs. Using the estimates of this study, 30,066,000 days of care for adults hospitalized in general acute care hospitals in 1981 could potentially have been eliminated and the care provided in other than an acute care setting. Not all these days can, of course, be easily eliminated. Allowing for (1) the 18% of nonacute admissions which are due to the unavailability of alternate care and the 24% of nonacute days occurring in justified admissions which occur for the same reason, (2) the 9% of nonacute days which are due to patient condition factors and/or patient education and (3) another 10% nonacute days which occur through inherent inefficiencies in the medical decision process, the data still suggest that over half the nonacute days could be eliminated without major structural changes to the health care system.

Savings beyond this level are possible if society could ensure the availability of alternate forms of care. The findings so far do not shed much light on the kinds of resource deficiencies which contribute to this problem. No definite relationship was found between nonacute use and nursing home beds per capita at the county level. It is not clear, therefore, whether the problem is due to the sheer unavailability of beds or is simply one of coordinating patients with available beds. Other issues, such as whether a subacute unit attached to the hospital promotes efficiency, also remain unexplored. With approximately one-fifth of all nonacute days attributable to the lack of alternate care, however, there is substantial potential for cost-savings to be made through providing the mechanisms and facilities for alternate care. Although no obvious correlations were found in this analysis, future analyses should concentrate on a more finely-tuned examination of the area and hospital factors specifically associated with the nonacute days that are due to the unavailability of alternate care.

Another set of questions still unanswered relates to the causal mechanisms through which nonacute utilization occurs. Such questions include (1) whether hospitals contribute to nonacute utilization at all and, if so, whether through the practice style of their physicians, through the unavailability of resources needed to ensure efficient provision of services or through financial incentives or organizational structures which promote or discourage overutilization; (2) whether physicians contribute to nonacute utilization and if so whether because of practice style differences associated with age, training, or regional professional standards or because of financial incentives which operate to promote unnecessary admissions; (3) whether patients contribute to nonacute utilization and if so whether a contributing factor is insurance coverage which fails to provide incentives for the control of utilization; and (4) whether location contributes to nonacute utilization above and beyond these other factors and, if so, whether because of resource unavailability or local expectations and practices . These questions are difficult to answer given the intercorrelations among the possible explanatory variables. More finely-tuned analysis of the day-level data set along with a causal modeling approach for both the admission analysis and day level analysis is needed to make a further contribution in this area. In this endeavor it is not the intention of the study to place blame on any one factor of the health care system. Knowing the circumstances under which nonacute utilization occurs, however, is central to understanding what actions must be taken in order to control it.

In this light, although the study contributed to an understanding of where nonacute utilization occurs in the stay, it was not possible to assess how it occurs. It is, therefore, not clear how much of nonacute utilization is attributable to inefficiencies that are inherent in the medical decision-making process, how much to hospital factors, how much to physician factors, and so forth. Studies performed on-site, concurrently with the patient stay may be the only way to ascertain this, as well as to identify the causal mechanisms through which inappropriate use operates.

In one sense the answers to these questions are irrelevant. With the introduction of the prospective payment system under Medicare and the possibility of widespread adoption of this system, society is coming closer to placing the burden of controlling nonacute utilization (at least for days of care) on the provider. Nevertheless, health care utilization is a complex phenomenon and the prospective payment system and other regulatory and competitive strategies may have a number of unintended consequences. Efforts to gain a better understanding of the processes which contribute to nonacute utilization can only help payors and providers alike to work together to provide the most cost-effective, high quality care possible.

To summarize, the present study estimated rates of nonacute admissions and days for adult patients in short-term acute care hospitals in 1981. A number of patient stay, hospital, and area factors associated with nonacute use were identified. In addition reviewer-supplied reasons for nonacute utilization were examined. This analysis, however, should be viewed as only the initial step in an ongoing program of research. As follow-up steps to the present research we see the following items as essential.

First, an independent evaluation of the Standardized Medreview Instrument should be conducted, comparing it with other utilization review instruments as well as independent physician judgment. Given the potential usefulness of the methodology developed for this project for utilization control by hospitals, Peer Review Organizations and other organizations, a thorough understanding of its properties can contribute not only to the interpretation of this study, but to practical cost containment efforts as well.

The analyses presented here do not begin to exhaust the potential of the data set. Further analyses could profitably be performed at a more refined level. More detailed analyses of the day level data set to identify factors associated with nonacute days by phase of stay may produce results where aggregate analysis did not. In addition, an examination of nonacute utilization arising through the unavailability of alternate care might also contribute to a greater understanding of this particular issue. Finally, given the complexity of the relationships among variables, a causal modeling approach to analysis might profitably be employed.

Analysis is unfortunately limited by the size of the sample obtained from each hospital and the lack of detailed data on various hospital organizational characteristics. If similar data collection efforts are conducted in the future, collecting more records per hospital and more precise information on hospital characteristics would certainly be worth considering. Similarly, collecting data on physician characteristics would considerably enhance the explanatory potential of the data. A real understanding of the processes and causal factors involved in nonacute utilization may, however, only be attainable through studies done on-site concurrently with the patient stay.

REFERENCES

Ahlmark, G. et al., "A Controlled Study of Early Discharge after Uncomplicated Myocardial Infarction," Acta Med Scandinavia, 206:87, 1979.

American College of Surgeons, General Surgery Criteria Sets for PSRO Screening, Chicago, 1976.

Anderson, J.G., "Demographic factors affecting health services utilization: a causal model," Medical Care, 11:104, 1973.

Averill, R.F. and McMahon, L.F., "A Cost Benefit Analysis of Continued Stay Certification," Medical Care, 15:158, 1977.

Berg, R.L., Browning, F.E., Crump, L., Wenkert, W., "Bed Utilization Studies for Community Planning, JAMA, 207:13, 2411-13, 1969.

Berki, S.E., "Hospital Economics," Lexington, MA: Lexington Books, 1972.

Berkman, B.G., and Rehr, H., "Early Social Service Case Finding for Hospitalized Patients: An Experiment," Social Services Review, 45:256, 1973.

Biggs, E.L., Kralewski, J.E., Brown, G.D., "A Comparison of Contract-Managed and Traditionally Managed Nonprofit Hospitals," Medical Care, 18:585, 1980.

Blumberg, M.S., "Regional Differences in Hospital Use Standardized by Reported Morbidity," Medical Care, 20:931, 1982.

Boaz, R.F., "Utilization Review and Containment of Hospital Utilization, Some Implications of Providing Care in the 'Most Appropriate Setting'", Medical Care, 17:315, 1979.

Borchardt, P.J., "Non Acute Profiles: Evaluation of Physicians' Nonacute Utilization of Hospital Resources," Quality Review Bulletin, November, 1981.

Brewer, W.R. and Freedman, M.A., "Causes and Implications of Variation in Hospital Utilization," Journal of Public Health Policy, 3:445, 1982.

Brian, E.W., "Government Control of Hospital Utilization - A California Experience," New England Journal of Medicine, 286:1340, 1972.

Brook, R.H. and Stevenson, R.L., "Effectiveness of Patient Care in An Emergency Room," New England Journal of Medicine, 283:904, 1970.

Brook, R.H., Williams, K.N., Rolph, J.E., "Use, Costs and Quality of Medical Services: Impact of the New Mexico Peer Review System," Annals of Internal Medicine, 89:256, 1978.

Browning, F.E., "The Record in Hospital Bed Utilization," In: Utilization Review: A Handbook for the Medical Staff. Medical Association, Department of Hospital and Medical Facilities, Chicago, 1965.

Bunker, J.F., "Surgical Manpower: A Comparison of Operations and Surgeons in the United States and in England and Wales," New England Journal of Medicine, 282:135, 1970.

Burford, R. and Averill, R.F., "The Relationship between Diagnostic Information Available at Admission and Discharge for Patients in One PSRO Setting," Medical Care, 17:369, 1979.

Calogero, M.A. and Hill, D.B., "Periodic Medical Review: The Impact in Pennsylvania," Medical Care, 20:85, 1982.

Chase, M., Laszio, C., Uyeno, D., "Modelling a Hospital Entrance System," Dimensions in Health Service, 57:16, 1980.

Chassin, M.R., "The Containment of Hospital Costs: A Strategic Assessment," Medical Care Supplement, 16:1, 1978.

Commerce Clearing House, 1982, "Medicare and Medicaid, Hospital Prospective Payment for Medicare," No. 374, Jan. 5, 1983, Chicago, Illinois.

Cordle, F. and Tyroler, H.A., "The Use of Hospital Medical Records for Epidemiologic Research. I. Differences in Hospital Utilization and In-Hospital Mortality by Age-Race-Sex-Place of Residence and Socioeconomic Status in a Defined Community Population," Medical Care, 12:596, 1974.

Coyne, J.S., "Hospital Performance in Multihospital Systems: A Comparative Study of System and Independent Hospitals," Health Services Research, 17:4, 1982.

Currie, C.T., Smith, R.G., Williamson, J., "Medical and Nursing Needs of Elderly Patients Admitted to Acute Medical Beds," Age and Aging, 8:49, 1979.

Davidson, S.M., Wacker, R.C., Klem, D.H., "Professional Standards Review Organizations: Critique," JAMA, 226:1106, 1973.

Detmer, D.E. and Tyson, T.J., "Regional Differences in Surgical Care Based Upon Uniform Physician and Hospital Discharge Abstract Data," Annals of Surgery, 187:166, 1978.

Dikewood Industries, Inc., "Assessment of the Factors Which Impact on the Accuracy of Concurrent Review Decisions," Washington, D.C., Department of Health, Education and Welfare. DHEW Publ. No. (HCFA/HSQB) 79-30000, 1979.

Dittman, D.A. and Magee, R.P., "Optimal Investigation Policies Under Selected PSRO Procedures," Medical Care, 18:1032, 1980.

Donabedian, A., "A Guide to Medical Care Administration, Vol.2, Medical Care Appraisal - Quality and Utilization" (American Public Health Association) 1969.

Donabedian, A., Methods for Deriving Criteria for Assessing the Quality of Medical Care," Medical Care Review, pp.653-698, Fall, 1980.

Donabedian, A., "The Quality of Medical Care," Science, 200(26): 856, 1978.

Drummond, M.F. and Mooney, G.H., "Part V - Assessing the Costs and Benefits of Treatment Alternatives," British Medical Journal, 285:1638, 1982.

Edwards, A.B., "Care Level and Timeliness Review - An Approach to Curbing Inappropriate Hospital Utilization," Topics in Health Care Financing, Vol. 7, No. 3:47, 1981.

Elliot, R.V., Kahn, K.A., Kaye, R., "Physicians Measure Up: A Study of 13 Surgical Procedures," JAMA, 245:595, 1981.

Eisenberg, J.M., and Rosoff, A.J., "Physician Responsibility For the Cost of Unnecessary Medical Services," New England Journal of Medicine, 299:76, 1978.

Escovitz, G.H. et al., "The Effects of Mandatory Quality Assurance: A Review of Hospital Medical Audit Processess," Medical Care, 16:941, 1978.

Federal Register (31 FR1 3424), Oct 18, 1966.

Feldstein, M.S., "Economic Analysis for Health Science Efficiency," Amsterdam: North-Holland Publishing Co., 1967.

Fitzpatrick, T.B., Riedel, D.C., Payne, B.C., "Appropriateness of Admission and Length of Stay," In: McNerney, W.J. et al., Hospital and Medical Economics, Chicago, Hospital Research and Educational Trust, 1962, Vol. 1.

Flashner, B.A. et al., "Professional Standards Review Organizations, Analysis of Their Development and Implementation Based on a Preliminary Review of the Hospital Admission and Surveillance Program in Illinois," JAMA, 223:1473, 1973.

Flood, B. et al., "The Relationship Between Intensity and Duration of Medical Services and Outcomes for Hospitalized Patients," Medical Care, 17:1088, 1979.

Fuhs, P.A., Martin, J.B., Hancock, W.M., "The Use of Length of Stay Distributions to Predict Hospital Discharges," Medical Care, 17:355, 1979.

Fulchiero, A. et al., "Can the PSRO's Be Cost Effective? A Study of the Effect of the Commonwealth Health Agencies Monitoring Program on the Length of Stay of Medical Patients in Massachusetts," New England Journal of Medicine, 299:574, 1978.

Gaffney, J.C. and Glandon, G.L., "The Physician's Use of the Hospital," Health Care Management Review, 7:49, 1982.

Gertman, P.M. and Bucher, B.M., "Inappropriate Hospital Bed Days and Their Relationship to Length of Stay Parameters," Unpublished Paper presented at the Medical Care Section, 99th Annual Meeting of the American Public Health Association, Minneapolis, October 11, 1969.

Gertman, P.M. et al., "Second Opinions For Elective Surgery: The Mandatory Medicaid Program in Massachusetts," New England Journal of Medicine, 302:1169, 1980.

Gertman, P.M. et al., "Utilization Review in the United States: Results from A 1976-1977 National Survey of Hospitals," Supplement to Medical Care, 17:1, 1979.

Gertman, P.M. and Restuccia, J.D., "The Appropriateness Evaluation Protocol: A Technique for Assessing Unnecessary Days of Hospital Care," Medical Care, 19:855, 1981.

Gibson, R.M. and Mueller, M., "National Health Expenditures, Fiscal Year 1976," Social Security Bulletin, 40:13, 1977.

Gibson, R.M. and Fisher, C.R., "National Health Expenditures, Fiscal Year 1977," Social Security Bulletin, 41:3, 1978.

Ginsberg, P.B. and Koretz, D.M., "Bed Availability and Hospital Utilization: Estimates of the 'Roemer Effect'," Health Care Financing Review, 5:87, 1983.

Glass, R.I. et al., "The 4 Score: An Index for Predicting A Patient's Non-medical Hospital Days," American Journal of Public Health, 67:751, 1977.

Goldberg, G.A. and Holloway, D.C., "Emphasizing 'level of care' over 'length of stay' in Hospital Utilization Review," Medical Care, 12:474, 1975.

Goldfarb, M,. Hornbrook, M., Rafferty, J., "Behavior of the Multiproduct Firm: A Model of the Nonprofit Hospital System," *Medical Care*, 18:185, 1980.

Goldstone, J. and Way, L.W., "The Use of Medical Audits in Surgical Education," *Surgery*, 84:25, 1978.

Gonnella, J.S., Miller, L.A., Smithline, H., "Identifying Patient Care Problems by Analyzing Critical Indicator Data," *Quality Review Bulletin*, p.16, September, 1980.

Gonnella, J.S. and Zeleznik, C., "Factors Involved in Comprehensive Patient Care Evaluation," *Medical Care*, 12:928, 1974.

Goran, M.J., "The Evolution of the PSRO Hospital Review System," *Medical Care Supplement*, 16:1, 1978.

Gornick, M. "Ten Years of Medicare," *Social Security Bulletin*, 4:3, 1976.

Gornick, M., "Trends and Regional Variations in Hospital Use Under Medicare," *Health Care Financing Review*, 3:41, 1982.

Grafe, W.R. et al., "The Elective Surgery Second Opinion Program," *Annals of Surgery*, 188:323, 1978.

Gray, S.P. and Steffy, W., *Hospital Cost Containment Through Productivity Management*, p.1-22, New York: Van Nostrand Reinhold Company, 1983.

Greenberg, S.M. and Rogers, T.F., "Study Uncovers 'Unneccessary' Long-Term Stays in a Short-Term Hospital," *Hospitals, J.A.H.A.*, 41:76, 1967.

Gurney, M. and Jewett, R.S., "Constructing Orthogonal Replications for Variance Estimation," *Journal of the Merican Statistical Association*, 70:819, 1975.

Hall, D. and Bytheway, B., "The Blocked Bed Definition of a Problem," *Social Science Medicine*, 16:1985, 1982.

Hall, D.J., Bytheway, W.R., Williams, W.O., "A Comment on 'The Reliability and Application of Clinical Judgment'," Letter to the editor, *Medical Care*, 15:527, 1977.

Hansen, M.H., Hurwitz, W.N., Madow, W.G., *Sample Survey Methods and Theory*, New York: John Wiley and Sons, Inc., 1953.

Harris, D.M., "An Elaboration of the Relationship Between General Hospital Bed Supply and General Hospital Utilization," *Journal of Health and Social Behavior*, 16:163, 1975.

Harris, D.M., "Effect of Population and Health Care Environment on Hospital Utilization," *Health Services Research*, 10:229, 1975.

Hastings, J.E.F. et al., "Prepaid Group Practice in Sault St. Marie," Ontario: Part I: Analysis of Utilization Records, Medical Care, 11:91, 1973.

Health Care Financing Administration, "Professional Standards Review Organizations, 1978 Program Evaluation," Baltimore, M.D., Health Care Financing Administration Research Report, USDHHS, HCFA, HHS Publication, HCFA Publ. No. 03003, 9-79 (1979).

Health Care Financing Administration, "Professional Standards Review Organizations, 1979 Program Evaluation, Health Care Financing Administration Research Report, USDHHS, HCFA, HHS Publication, HCFA Publ. No. 03041, 5-80 (1980).

Health Policy Institute, Graduate School of Public Health, University of Pittsburgh, "The Pattern of Utilization of Inpatient Hospital Services in Southwestern Pennsylvania, HPI Policy Series # 1, November, 1980.

Health Resources Administration, The Area Resource File, DHHS Publication No. (HRA) 81-9, November, 1980.

Health Services Administration, Office of Planning, Evaluation and Legislation, "PSRO: An Evaluation of Professional Standards Review Organizations. Exeuctive Summary," Professional Standards Review Organizations: Program Evaluation. Report No. OPEL 77-12, 1977.

Heaney, C.T. and Riedel, D.C., "From Indemnity to Full Coverage: Changes in Hospital Utilization," Blue Cross Reports, Research Series 5, Blue Cross Assn., Chicago, October, 1970.

Hersey, F. and O'Neill P., "Reducing Post-Operative Hospital Stays," Dimensions in Health Service, 57:22, 1980.

Hess, I., Riedel, D.C., Fitzpatrick, T.B., Probability Sampling of Hospitals and Patients, Ann Arbor: Health Administration Press, 1975.

Holloway, D.C., Holton, J.P., Goldberg, G.A., Restuccia, J.D., "Development of Hospital Levels of Care Criteria," Health Care Management Review, 1:61, 1976.

Horn, S.D., Roveti, G.C., Kreitzer, S.L., "Length of Stay Variations: A Focused Review," Quality Review Bulletin, Febuary, 1980.

Ingman, S.R. and Claus, M.L., "Preadmission Assessments for Geriatric Patients," Quality Review Bulletin, September, 1981.

Johnson, E.A., "Organizational Theory Implications in Increasing Revenues: the Multihospital System," Topics in Health Care Financing, 8:84, 1981.

Kelman, H. et al., "Monitoring Patient Care," Medical Care, 7:1, 1969.

Kessner, D.M., Kalk, C.E., Singer, J., "Assessing Health Quality - The Case for Tracers," New England Journal of Medicine, 288:189, 1973.

Knaus, W.A. et al., "APACHE - Acute Physiology and Chronic Health Evaluation: A Physiologically Based Classification System," Critical Care Medicine, 9:591, 1981.

Koran, L.M., "The Reliability of Clinical Methods, Data and Judgements," Part One, New England Journal of Medicine, 293:642, 1975.

Koran, L.M., "The Reliability of Clinical Methods, Data and Judgements," Part Two, New England Journal of Medicine, 293:695, 1975.

Kurylo, L.L., "Measuring Inappropriate Utilization," Health and Hospital Services Administration, Winter, 1976.

Lave, J.R. and Leinhardt, S., "The Cost and Length of a Hospital Stay," Inquiry, 13:327, 1976.

Leske, M.C., Sorensen, A.A., Zimmer, J.G., "Discrepancies Between Admission and Discharge Diagnoses in a University Hospital," Medical Care, 16:740, 1978.

Levit, K.R., "Personal Health Care Expenditures by State, Selected Years 1966-1978," Health Care Financing Review, 4:1, 1982.

Lewis, C.E., "Variations in the Incidence of Surgery," New England Journal of Medicine, 281:880, 1969.

Longest, B.B., "An External Dependence Perspective of Organizational Strategy and Structure: The Community Hospital Case," Hospital and Health Services Administration, 26:50, 1981.

Longest, B.B., "Relationships Between Coordination, Efficiency, and Quality of Care in General Hospitals," Hospital Administration, 19:65, 1974.

Lubitz, J. and Deacon, R., "The Rise in the Incidence of Hospitalizations for the Aged, 1967 to 1979," Health Care Financing Review, 3:21, 1982.

Luft, H.W., Health Maintenance Organizations, pp.139-150, 358-359. New York: Wiley, 1980.

Luft, H.W., "How Do Health Maintenance Organizations Achieve Their 'Savings'?" New England Journal of Medicine, 298:1336, 1978.

Luke, R.D. and Thomson, M.A., "Utilization of Within-Hospital Services: A Study of the Effects of Two Forms of Group Practice," Medical Care, 18:219, 1980.

Martin, S.G. et al., "Impact of Mandatory Second-Opinion Program on Medicaid Surgery Rates," Medical Care, 20:21, 1982.

McCarthy, E.G and Widmer, G.W., "Effects of Screening by Consultants on Recommended Elective Surgical Procedures," New England Journal of Medicine, 291:1331, 1974.

McCarthy, E.G. and Finkel, M.L., "Second Consultant Opinion for Elective Surgery Programs: Outcome Status Over Time," Medical Care, 16:984, 1978.

McClain, J.O., "Physician Confidence and Reliability in Utilization Review," Medical Care, 10:463, 1972.

McClain, J.O. and Riedel, D.C., "Screening For Utilization Review: On the Use of Explicit Criteria and Non-Physicians in Case Selection," American Journal of Public Health, 63:247, 1973.

McClure, W., "Toward Development and Application of a Qualitative Theory of Hospital Utilization," Inquiry, 19:117, 1982.

McCracken, G., "Determinants of Hospital Use: An Examination of County Variations in Iowa," Mimeograph, Des Moines: Iowa Health Systems Agency, 1980.

McKay-Dee Hospital Center. Multilevel Care. 3939 Harrison Boulevard, Ogden, Utah 84402, updated.

McKinlay, J.B., "Some Approaches and Problems in the Study of the Use of Services - An Overview," Journal of Health and Social Behavior, 13:115, 1972.

Mechanic, D., "Approaches to Controlling the Costs of Medical Care: Short-Range and Long-Range Alternatives," New England Journal of Medicine, 298:249, 1978.

The Medicare and Medicaid Data Book, 1981, "Health Care Financing Program Statistics," Department of Health and Human Services, HCFA, Pub. No. 03128, 1982.

Morehead, M.A., Donaldson, R., et al., "A Study of the Quality of Hospital Care Secured by a Sample of Teamsters' Family Members in New York City," New York: Columbia University School of Public Health and Administrative Medicine, 1964.

Mott, F.D., Hastings, J.E.F., Barclay, A.T., "Prepaid Group Practice in Sault Ste. Marie, Ontario: Part II: Evidence from the Household Survey," Medical Care 11:173, 1973.

Nassau County Medical Society, Voluntary Insurance Committee: Pilot Study of Hospital Use in Nassau County, Garden City, New York, November, 1963.

National Center for Health Services Research, "Controlling the Cost of Health Care," Washington, D.C., U.S. Department of Health, Education and Welfare. DHEW Publ. No. (HRA) 77-3182, 1977.

National Center for Health Statistics: Health in the United States: 1975. Washington, D.C., U.S. Department of Health, Education and Welfare. DHEW Publ. No. (HRA) 76-1233, 1976.

National Center for Health Statistics: "Pseudoreplication Further Evaluation and Application of the Balanced Half-Sample Technique," *Vital and Health Statistics*. Series 2-Number 31 (PB80-126188) Public *Health Service, Washington*, D.C. January, 1969.

National Center for Health Statistics: "Replication, an Approach to the Analysis of Data from Complex Surveys," *Vital and Health Statistics*. Series 2-Number 14 (PHS) 79-1269 Public *Health Service, Washington*, D.C., 1966.

Newhouse, J.P., "Insurance Benefits, Out-of-Pocket Payments, and the Demand for Medical Care: A Review of the Literature," *Hlth. Med. Care. Serv.*, 1(4):1, 1978a.

Newhouse, J.P., "The Structure of Health Insurance and the Erosion of Competition in the Medical Marketplace," in W. Greenberg (ed.), *Competition in the Health Care Sector: Past, Present, and Future*, Germantown, M.D.: Aspen Systems Corporation, 1978b.

Newhouse, J.P., Manning, W.G., Morris, C.N., Orr, L.L., Duan, N., Keeler, E.B., Leibowitz, A., Marquis, K.H., Marquis, M.S., Phelps, C.E., Brook, R.H., "Some Interim Results from a Controlled Trial of Cost Sharing in Health Insurance," Rand Corporation, Santa Monica, CA, 1982.

Nobrega, F.T., Krishan, I., Smoldt, R.K., Davis, C.S., Abbott, J.A., Mohler, E.G., McClure, W., "Hospital Use in a Fee-for-Service System," *JAMA*, 247:806, 1982.

Payne, B.C., "Continued Evolution of a System of Medical Care Appraisal," *JAMA*, 201:126, 1967.

Payne, B.C., *Utilization Review: A Handbook for the Medical Staff*, American Medical Association, Chicago, 1965.

Paulyh,1 M. and Redisch, M., "The Not-for-Profit Hospital as a Physician's Cooperative," *American Economic Review*, 63:87, 1973.

Querido, A., "Indirect Evaluation of a Randomized Sample of the General Practitioner's Work" (1954), In: *The Efficiency of Medical Care*, Leiden: H.E. Stenfert Kroese, NV., 1963.

Rafferty, J., "Enfranchisement and Rationing: Effects of Medicare on Discretionary Hospital Use," Health Services Research, 10:51, 1975.

Restuccia, J.D., "The Effect of Concurrent Feedback in Reducing Inappropriate Hospital Utilization," Medical Care, 20:46, 1982.

Restuccia, J.D. and Holloway, D.C., "Barriers to Appropriate Utilization of an Acute Facility," Medical Care, 14:559, 1976.

Richardson, F.M., "Peer Review of Medical Care," Medical Care, 10:29, 1972.

Ro, K.K., "Patient Characteristics, Hospital Characteristics and Hospital Use," Medical Care, 7:295, 1969.

Roemer, M., and Shain, M., Hospital Utilization Under Insurance. Chicago: American Hospital Association, 1959.

Roemer, M.I. and Shonick, W., "HMO Performance: The Recent Evidence," Milbank Memorial Fund Quarterly, 5:271, 1973.

Roos, N.P., Roos, L.L., Henteleff, P.D., "Elective Surgical Rates: Do High Rates Mean Lower Standards?" New England Journal of Medicine, 297:360, 1977.

Rosenfeld, L.S., Goldmann, F., Kaprio, L.A., "Reasons for Prolonged Hospital Stay; A Study of Need for Hospital Care," Journal of Chronic Disease, 6:141, 1957.

Rosenthal, G.D. "Factors Affecting the Utilization of Short-Term General Hospitals," American Journal of Public Health, 55:1734, 1965.

Rosser, R.M., "The Reliability and Application of Clinical Judgement in Evaluating the Use of Hospital Beds," Medical Care, 14:39, 1976.

Rothberg, D.L. and Gertman, P.M., "The Use of Attitudinal Data for Public Policy: The Case of Unnecessary Hospital Use," Medical Care, 19:47, 1981.

Rothberg, D.L., Pinto, J., Gertman, P.M., "Medicare Utilization in the United States: PSRO and Regional Impacts," Health Services Research, 15:249, 1980.

Rubin, L., "Comprehensive Quality Assurance Systems," American Group Practice: Alexandria, Va, 1975.

Rubin, S.G. and Davies, G.H., "Bed Blocking By Elderly Patients in General Hospital Wards," Age and Aging, 4:142, 1975.

Ruchlin, H.S., Finkel, M.L., McCarthy, E.G., "The Efficacy of Second Opinion Consultation Programs: A Cost-Benefit Perspective," Medical Care, 20:3, 1982.

Rushing, W., "Differences in Profit and Nonprofit Organizations: A Study of Effectiveness and Efficiency in General Short-Stay Hospitals," Administrative Science Quarterly, 19:476, 1974.

Sammons, J.H., "Statement by the American Medical Association on Unnecessary Surgery," J Med Assoc Ga, 66:633, 1977.

Sanazaro, P.J. and Worth, R.M., "Concurrent Quality Assurance in Hospital Care," New England Journal of Medicine, 298:1171, 1978.

Sayetta, R.B., "Critique of an earlier study of the Sacramento Medical Care Foundations' Certified Hospital Admission Program (CHAP)," Medical Care, 14:80, 1976.

Schwartz, J., "Management Rounds: Inflation Fueled Fourth Quarter Climb in '81," Hospitals, 56:36, 1982.

Scitovsky, A.A. and McCall, N., "Use of Hospital Services Under Two Prepaid Plans," Medical Care, 18:30, 1980.

Scott, W.R., Flood, A.B., Ewy, W., "Organizational Determinants of Services, Quality and Cost of Care in Hospitals," Milbank Memorial Fund Quarterly, 57:234, 1979.

Selvin, S., "Three Statistical Models for Estimating Length of Stay," Health Services Research, 12:322, 1977.

Shapiro, E. and Roos, N.D., "The Geriatric Long-Stay Hospital Patient: A Canadian Case Study," Journal of Health Politics, Policy and Law, 6:49, 1981.

Sherman, H. and Flatley, M., "Dissecting the Hospital Stay: A Method for Studying Patient Staging in Hospitals," Medical Care, 18:715, 1980.

Shortell, S.M., "Factors Associated with the Utilization of Health Services," In: Introduction to Health Services, Williams S.J. and Torrens, P.R., eds., pp.48-90, New York: Wiley and Sons, 1980.

Shortell, S.M., "Hospital Medical Staff Organization: Structure, Process and Outcome," Hospital Administration, 2:96, 1974.

Shortell, S.M., Becker, S.W., Neuhauser, D., "The Effects of Management Practices on Hospital Efficiency and Quality of Care," In: Organizational Research in Hospitals, Shortell, S.M. and Brown, M., eds., Publ. Blue Cross Association, 1976.

Sieverts, S., "The Assurance of Appropriate Medical Care: The Role of Blue Cross and Blue Shield," Bulletin of the New York Academy of Medicine, 58:49, 1982.

Slee, V.N., "PSRO and the Hospital's Quality Control," Annals of Internal Medicine, 81:97, 1974.

Sloan, F.A. and Steinwald, B., "Effects of Regulation on Hospital Costs and Input Use," Journal of Law and Economics, 23:81, 1980.

Smits, H.L., "The PSRO in Perspective," New England Journal of Medicine, 305:253, 1981.

Stockwell, H. and Vayda, E., "Variations in Surgery in Ontario," Medical Care, 17:390, 1979.

Stuart, B. and Stockton, R., "Control Over the Utilization of Medical Services," Health and Society, Summer: 341, 1973.

Studnicki, J. and Honemann, D., "Analyzing Inpatient Hospital Duration and Intensity: A Methodology," Quality Review Bulletin, September: 15, 1982.

SysteMetrics, Inc., "Identification of the Factors Involved in Differential Hospital Utilization Patterns", Field Staff Training Manual, Santa Barbara, CA, 1982.

Taylor, L.D. and Brooks, G.W., "A Screening Program to Reduce Admissions to a State Hospital," Hospital and Community Psychiatry, 31:59, 1980.

Trussell, R.E., Morehead, M.A., Erlich, J., The Quantity, Quality and Costs of Medical and Hospital Care Secured by a Sample of Teamster Families in the New York Area. New York: Columbia University Press for the Teamsters' Joint Council No. 16 and Management Hospitalization Trust Fund, 1962.

U.S. Bureau of the Census, "Statistical Abstract of the United States 1982-1983," (103rd edition), Washington, D.C., 1982.

Vayda, E. and Anderson, G.D., "Comparison of Provincial Surgical Rates in 1968," Canadian Journal of Surgery, 18:18, 1968.

Veeder, N.W., "Health Services Utilization Models for Human Services Planning," Journal of the American Institute of Planners, 41:101, 1973.

Wallace, S.G. and Anderson, A.D., "Imprisonment of Patients in Course of Rehabilitation," Archives of Physical Medicine and Rehabilitation, 59:424, 1978.

Weil, T.P., "Do More Physicians Generate More Hospital Utilization?" Hospitals, 55:70, 1981.

Wenkert, W., Hill, J.G., Berg, R.L., "Concepts and Methodology in Planning Patient Care Services," Medical Care, 7:4, 327-31, 1969.

Wennberg, J.E., Barnes, B.A., Zubkoff, M., "Professional Uncertainty and the Problem of Supplier-induced Demand," Social Science Medicine, 16:811, 1982.

Wennberg, J.E. and Fowler F.J, "A Test of Consumer Contributions to Small Area Variations in Health Care Delivery," J. Maine Med Assoc, 68:275, 1977.

Wennberg, J.E. and Gittelsohn, A., "Health Care Delivery in Maine I: Patterns of Use of Common Surgical Procedures," J. Maine Med Assoc, 66(5):123-130 and 149, 1975.

Wennberg, J.E., Gittelsohn, A. and Soule, D., "Health Care Delivery in Maine II: Conditions Explaining Hospital Admissions," J. Maine Med Assoc, 66(10):255-261 and 269, 1975.

Wennberg, J. and Gittelsohn, A., "Small Area Variations in Health Care Delivery," Science, 182:1102, 1973.

Wennberg, J.E. and Gittelsohn, A.M., "Variations in Medical Care Among Small Areas," Scientific American, 246(4): 120-134, April 1982.

Westphal, M., Frazier, E., McClinton, M., "Changes in Average Length of Stay and Average Charges Generated Following Institution of PSRO Review," Health Services Research, 14:253, 1979.

Wilensky, G.R. and Rossiter, L.F., "The Relative Importance of Physician-induced Demand in the Demand for Medical Care," Milbank Memorial Fund Quarterly, 61:252, 1983.

Williamson, J.W., Improving Medical Practice in Health Care: A Bibliographic Guide to Information Management in Quality Assurance and Continuing Education. Ballinger Publishing Co: Cambridge, MA 1979.

Wylie, C.M. and Flashner, B.A., "Reassessing the Early Effect of Concurrent Review on Length of Stay," Medical Care, 17:607, 1979.

Wymelenberg, S., "Aggressive Utilization Review Shortens Stay, Lowers Costs," Hospitals, 52:75, 1978.

Zimmer, J.G., "An Evaluation of Observer Variability in a Hospital Bed Utilization Study," Medical Care, 5:221, 1967.

Zimmer, J.G. and Groomes, E.W., "An Observer Reliability Study of Physicians' and Nurses' Decisions in Utilization Review of Chronic-care Facilities," Medical Care, 7:14, 1969.

Zimmer, J.G., "Length of Stay and Hospital Bed Misutilization," Medical Care, 7:453, 1974.

Zuckerman, H.S., "Multi-Institutional Systems: Promise and Performance," Inquiry, 16:291, 1979.

APPENDIX A

TECHNICAL ADVISORY PANEL and MEDICAL CONSULTANT PANEL MEMBERS

APPENDIX A:

TECHNICAL ADVISORY PANEL (TAP) MEMBERS

- Joseph S. Gonnella, M.D.
  Associate Dean and Director of Academic Programs
  Jefferson Medical College
  Philadelphia, PA

- Michael D. Intriligator, Ph.D.
  Department of Economics
  University of California at Los Angeles

- Mary Etta Mills, R.N., Sc.D.
  Director of Quality Assurance, Research and Education
  Department of Nursing
  University of Maryland

- John Emerson Ware, Jr., Ph.D.
  Social Science Department and Health Program
  The Rand Corporation
  Los Angeles, CA

- Elizabeth M. Baldwin, A.C.S.W.
  Social Work Consultant
  Greater Southeast Community Hospital
  Washington, D.C.

- Beaufort B. Longest, Jr., Ph.D.
  Director, Health Policy Institute
  School of Public Health
  University of Pittsburgh

- Dale N. Schumacher, M.D.
  President, Rockburn Institute
  Elkridge, MD

- Bruce M. Bucher, M.D.
  Health Services Researcher and Practicing Physician
  Tappahannuck, VA

- Peter Borchardt, M.B.A.
  DelMarva Foundation for Medical Care, Inc.
  Easton, MD

- Geoffrey Gibson, Ph.D.
  Professor of Health Care Administration
  Baruch College
  Manhattan, NY

- Faye Brown, R.R.A.
  Former President
  American Medical Records Assn
  St. Helena, CA

SAMPLING CONSULTANT:

- Joseph Steinberg
  Survey Design, Inc.
  1011 Roswell Drive
  Silver Spring, MD

APPENDIX A (continued)

LIST OF MEDICAL CONSULTANT PANEL MEMBERS

- George Fisher, M.D.
  Internal Medicine
  Pennsylvania Hospital
  Philadelphia, PA

- Joseph Gonnella, Ph.D.
  Associate Dean & Director
  of Academic Programs
  Jefferson Medical College
  Philadelphia, PA

- Frederick D. Burg, M.D.
  Associate Dean for Academic Programs
  School of Medicine
  University of Pennsylvania
  Philadelphia, PA

- Richard K. Spence, M.D.
  General Vascular Surgeon
  Cooper Surgical Association
  Camden, NJ

The following individuals also provided advice and consultation to the project:

Paul Gertman, M.D.
Boston University Medical Center
Health Care Research
Boston, MA

Jean Murphy
Director of Review Systems
Delmarva Foundation for Medical Care
Easton, MD

Lyla Hernandez
Executive Director
American Association of PSROs
Washington, D.C.

Thomas Fitzpatrick
Graduate School of Public Health
University of Pittsburgh
Pittsburgh, PA

Kirsten Kingdon
Executive Director
Michigan Area VII PSRO
Ann Arbor, MI

Thomas Terrell, Ph.D.
Becky Zuckerelli
Akron City Hospital
Akron, OH

Wanda Young, Sc.D.
Blue Cross of Western Pennsylvania
Pittsburgh, PA

Sheila Richard
National Health Standards and
Quality Information Clearinghouse
Kensington, MD

Harlan Bennett
Executive Director
Santa Clara Valley PSRO
San Jose, CA

George Goldberg, Ph.D.
Rand Corporation
Santa Monica, CA

Greg Banks
Vice-President
Interqual, Inc.
Bethesda, MD

Joanne Lamprey, RN
Interqual
Bethesda, MD

*U.S. GOVERNMENT PRINTING OFFICE: 421-100/15226

CMS LIBRARY
3 8095 00012234 7

U.S. DEPARTMENT OF
HEALTH AND HUMAN SERVICES
HEALTH CARE FINANCING ADMINISTRATION
OFFICE OF RESEARCH AND DEMONSTRATIONS
6325 SECURITY BLVD.
BALTIMORE, MARYLAND 21207

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF H.H.S.
HHS-392

SPECIAL FOURTH CLASS
BOOK RATE

U.S. MAIL ®

U.S. Department of Health and Human Services
Health Care Financing Administration
Office of Research and Demonstrations
HCFA Pub. No. 03182 September 1984